

## **General Disclaimer**

### **One or more of the Following Statements may affect this Document**

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.
- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.
- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.
- This document is paginated as submitted by the original source.
- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

CR-152157

MIT

AN ANALYSIS OF SHORT HAUL  
AIR PASSENGER DEMAND

(NASA-CR-152157) AN ANALYSIS OF SHORT HAUL  
AIR PASSENGER DEMAND, VOLUME 2 Final Report  
(Massachusetts Inst. of Tech.) 132 p HC  
A07/MF A01

N79-22063

CSCL 05C

G3/03 Unclas  
25583

DEPARTMENT  
OF  
AERONAUTICS  
&  
ASTRONAUTICS

TERRY P. BLUMER  
&  
WILLIAM M. SWAN

1978



FLIGHT TRANSPORTATION  
LABORATORY  
Cambridge, Mass. 02139

NASA CR 152157

AN ANALYSIS OF SHORT HAUL AIR PASSENGER DEMAND  
Volume II

FLIGHT TRANSPORTATION LABORATORY  
DEPARTMENT OF AERONAUTICS & ASTRONAUTICS  
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

1978

The research was supported by the Research Aircraft Projects Office of Ames Research Center, National Aeronautics and Space Administration under NASA Grant No. NSG-2129. The research group wishes to acknowledge the technical assistance provided by Louis J. Williams and Mark H. Waters of the Ames Research Center.

## TABLE OF CONTENTS

List of Tables .....	i
List of Figures .....	i
List of Symbols-PART I .....	ii
List of Symbols-PART II .....	iv
Abstract .....	1
Section 1. Introduction .....	2
Section 2. Model Specification and Sequential Results .....	7
2.1 Base Model .....	7
2.1.1 Gravity Term .....	8
2.1.2 Air Impedance .....	9
2.1.3 Level of Interest .....	10
2.1.4 Empirical Results of the Base Model .....	12
2.2 Mode Sensitive Model .....	14
2.2.1 Proxy Specifications .....	14
2.2.2 Air Share Specifications .....	17
2.2.3 Proxy Specifications Results .....	18
2.2.4 Air Share Specifications Results .....	20
2.2.5 Summary .....	20
2.3 Destination Sensitive Models .....	21
2.3.1 Model Specifications .....	23
2.3.2 Statistical Results .....	25
2.4 Mode-Destination Sensitive Model .....	27
2.4.1 Model Specifications .....	27
2.4.2 Results of the Combined Model .....	28
2.4.3 Statistical Verification for Linearity, Heteroscedascity, Normality, and Pooled Data .....	29
Section 3. Conclusions .....	34

## Table of Contents (CONT.)

### Part II.

Section 1.	Introduction .....	39 <sup>1</sup>
Section 2.	Traditional Models .....	41
Section 3.	A New Formula .....	46
3.1	Travel Generation .....	46
3.2	Alternative Destinations .....	47
3.3	Modal Split .....	50
3.4	Price/Consumption .....	52
Section 4.	Mathematical Formulation .....	54
4.1	Travel Generation .....	54
4.2	Alternative Destinations .....	54
4.3	Modal Split .....	55
4.4	Price Consumption .....	56
4.5	Summary .....	56
Section 5.	Calibration .....	62
Section 6.	Conclusions .....	66
Section 7.	Future Work .....	69
REFERENCES		76
APPENDIX A	Sources of Data .....	80
APPENDIX B	Empirical Support for the Short Haul Demand Model .....	84
APPENDIX C	List of Tables Pertaining to Model Results .....	107

## LIST OF TABLES

## PART II

Table 5.1	Data .....	63
Table 5.2	Variation in Parameters .....	64
Table 5.3	Tests for Robust Estimates .....	65

## LIST OF FIGURES

## PART I

Figure 1.1	Development of the Short Haul Demand Model .....	6
Figure 2.1	Derivation of Level of Interest .....	11
Figure 2.2	Effect of a Competing Destination on City-Pair Traffic .....	22

## PART II

Figure 3.1	Distribution of Trips vs. Distance for Constant Population Density, Intervening Opportunities Formulation .....	49
Figure 3.2	Conceptual Representation of Modal Split by Value of Time .....	51
Figure 7.1a	Modal split form 1968 Northeast Corridor Data ...	70
Figure 7.1b	Modal split form 1968 Northeast Corridor Data ...	70
Figure 7.2	Value of time distribution as indicated by data	71
Figure 7.3	Possible Value of Time Distribution .....	72

## List of Symbols-PART I

$a_1$  = Exponent of Distance - given level of interest effect only.

$AP_{ij}$  = Proxy for alternate destinations.

$b_0, b_1, b_2, b_3, b_4$  = Constants determined by regression.

$C_1, C_2, C_3$  = Constants determined by regression.

$CP_{ij}$  = Proxy for level of intermodal competition.

$D_{ij}$  = Distance between cities i and j.

$DA_{ij}$  = Difference between air and auto impedances.

$f_{ij}$  = Share of travelers using the air mode.

$F_{ij1}$  = Fraction of total travelers moving between cities i and j  
(Source-Two specification)

$F_{ij2}$  = Fraction of total travelers moving between cities i and j  
(Source-Two specification)

$G_{ij}$  = Travel attraction between cities i and j.

$GH_{ij1}$  = Total air travelers generated from cities i and j.  
(Source-One specification)

$GH_{ij2}$  = Total air travelers generated from cities i and j.  
(Source-Two specification)

$H_{ij}$  = Harmonic mean of rail and auto impedances.

$\bar{I}_{ij}$  = Harmonic mean of air, rail and auto impedances.

$I_{ija}$  = Disutility of air travel between cities i and j, also called air impedance.

$I_{ijr}$  = Rail impedance.



## List of Symbols-PART I

$I_{iju}$  = Auto impedance

$LOI$  = Level of interest

$M_i$  = Effective buying income of city i.

$M_j$  = Effective buying income of city j.

$S_i$  = Total attraction of destinations for travelers from city i.

$S_j$  = Total attraction of destinations for travelers from city j.

$T_{ij}$  = Total number of origin-destination air travelers between cities i and j.

$\theta$  = Exponent on generation terms, as generation may not be directly proportional to the source strength.

$V$  = Inverse of value of time.

## LIST OF SYMBOLS-PART II

$D_{ij}$	=	Total travel between cities i and j
$d_{ij}$	=	Distance between cities i and j
$P_i$	=	Population of city i
$P_j$	=	Population of city j
$K_m$	=	Conductance for mode m
$C_m$	=	Cost of using mode m
$T_m$	=	Travel time by using mode m
$A_m$	=	Other characteristics of mode m
$F_n$	=	Fraction of traffic allocated to mode n
$\Delta$	=	Probability of any one person satisfying the trip purpose
$\ell_{ij}$	=	Population within a circle with a radius $d_{ij}$ of city i, including the population of city i itself
$\Phi(v)$	=	Probability density function for value of time v.
$D_{ijn}$	=	Demand between cities i and j on mode n
$\alpha$	=	Trip consumption elasticity with respect to total of time and money cost
$\beta$	=	elasticity with respect to income
$V_{nm}$	=	Dollars per hour at which auto and air modes are equally expensive
PAX	=	1973 average daily one-way origin-destination passengers
$V_{avg}$	=	Average value of time among the traveling public assuming Gaussian distribution of such values
$V_\sigma$	=	Standard deviation of the Gaussian distribution of values of time
k	=	Scale factor for all demands
I	=	Average annual per capita income

## ABSTRACT

In Part 1 of this report, several demand models for short haul air travel are proposed and calibrated on pooled data. The models are designed to predict demand and analyze some of the motivating phenomena behind demand generation. In particular, an attempt is made to include the effects of competing modes and of alternate destinations.

The results support three conclusions: first, the auto mode is the air mode's major competitor; second, trip time is an overriding factor in intermodal competition, with air fare at its present level appearing unimportant to the typical short haul air traveler; and finally, distance appears to underly several demand generating phenomena, and therefore must be considered very carefully in any intercity demand model. It may be the cause of the wide range of fare elasticities reported by researchers over the past 15 years.

Part 2 of this report extends the work discussed in Part 1. A new behavioral demand model is proposed and calibrated. It combines the travel generating effects of income and population, the effects of modal split, the sensitivity of travel to price and time, and the effect of alternative destinations satisfying the trip purpose. This new behavioral model appears to be as accurate as the models developed in Part 1, and also offers a number of new promising directions for further research in short haul demand analysis.

PART I

An Analysis of Short Haul Air Passenger Demand

by

Terry P. Blumer

## SECTION 1

### INTRODUCTION

Among the numerous demand models that have appeared in the air transportation literature in recent years, some have explored new structural forms while others have examined special segments of the travel market (by region, passenger type, and so on). The models range from those which analyzed short haul demand within a given geographic area. No model to date, however, has adequately investigated the short haul market for a representative cross section of U.S. city-pairs.

The model developed and calibrated in this report is an attempt to close this gap. It is designed to forecast demand and analyze passenger behavior in U.S. domestic, short haul, air markets. In particular, the model captures the effect of intermodal competition and of alternate destinations upon the demand for air travel.

An extensive search was performed for an appropriate model structure. Of the many models investigated, the choice or mode split models seemed to offer the greatest potential because they inherently capture the effects of intermodal competition. These models use two equations: first, one called the trip generation equation which estimates total travelers by all modes and usually employs a gravity formulation; and second, one called the split equation which calculates the market share of each mode. Specification of the split equation creates the greatest interest and controversy among modelers, who attempt to use the attributes of the various modes (travel time, fare, frequency, etc.) to produce a reasonable traffic split. The

functional forms of the traditional choice models are the following:

(1) the trip generation equation where total travelers = function of (travel attraction, travel impedance, ...); and the split equation where the fraction of travelers using mode  $m$  = function of (attributes of mode  $m$  relative to attributes of other modes).

Typical intercity choice models have been developed by Quandt and Baumol, a Berkeley research team, and United Technologies Research Center (UTRC) (15, 20, 33). The research group at Berkeley developed a sophisticated split equation from a model which assumes that the coefficients of the mode attributes are random variables. This model's novel structure was tested on a set of city pairs drawn from California. Quandt and Baumol's model is called the Abstract Mode Model because the mode attributes such as travel time are weighted equally regardless of the mode. For example, an hour in an auto is equivalent to an hour on a train. Again the structure is novel and very complex, and tested on data drawn from a limited area (the Northeast Corridor).

UTRC developed a Probit model based upon an explicit transportation disutility function.\* Again the model is unique and quite sophisticated. It was calibrated on data from the National Transportation Survey, which incidentally had most of its significant city pairs drawn from California and the Northeast Corridor.

Although the choice models handle intermodal competition well, they do

---

\* The probit probability model is associated with the cumulative normal probability function. For further details see D.J. Finney, Probit Analysis, 2d ed. (Cambridge, England: Cambridge University Press, 1964), and James Tobin, "The Application of Multivariate Probit Analysis to Economic Survey Data" (Cowles Foundation Discussion Paper 1, Yale University, 1955).

contain certain well known difficulties. The choice models are still in the exploratory stage and as the variety of split equations indicate, the correct specification remains a major obstacle. Choice models require disaggregate data (by mode) which is not available at the national level. Consequently the choice models have been applied only regionally and in urban studies. The National Transportation Survey claims to be national in scope, but as UTRC discovered, it is strongly biased to the two coasts. Therefore mode choice models appear to be impractical for this short haul air model.

A model structure often used for intercity demand studies is the log-linear specification. While it has enjoyed only moderate success in past studies, it does have the advantage of yielding elasticities of the independent variables directly through the coefficients. It also has the advantage of requiring no disaggregated trip data by mode. It does not, however, inherently account for intermodal competition.

Typical log-linear models have been developed by Stuart, Alcala, and an M.I.T. research group (29, 1, 17). Stuart explored various specifications of the gravity formulation, and in fact most single equation, city-pair models depend upon a gravitational formulation. This formulation though is too simple to adequately explain passenger generation. Alcala built a model similar to Stuart's but added a second term, air fare, in order to improve the model and expand the description of the generation process. The M.I.T. research group created a new variable, called level of service, which describes the expected trip time based upon a behavioral model. This, together with a gravity term and air fare produces a very simple behavioral

model, but still not so detailed as the choice models which also include the effects of competing modes.

The trade-off among model structures was very simple. The choice models were behavioral in origin and therefore good for analysis but due to lack of trip data could not be calibrated reliably, while the log-linear models can be calibrated easily but typically do not have behavioral origins. Since a model that cannot be calibrated is impractical, no matter how sophisticated the specification, the choice specifications were set aside in favor of the log-linear specification.

This paper also considers the attraction of alternate destinations. It is assumed that a traveler divides his trips among the destinations available. If only a few destinations exist, then each may be visited often, but if many exist then each may be visited infrequently. This means the level of traffic between two cities depends upon the number of other cities in the surrounding area as well as the gravitational attraction, level of air service, intermodal competition, and air fare. The effect seems reasonable and therefore is tested in the short haul model.

Our short haul model is developed in a logical, step-by-step manner. First a base model is specified which resembles the simple log-linear models mentioned above, containing a gravity term and a level of air service term. Then intermodal competition is added to the base model and its effect investigated. Subsequently an alternate destination term is incorporated in the base model instead of the intermodal competition term.\* Finally, both intermodal competition and alternate destinations are included in the model

---

\* Mode-sensitive and destination-sensitive models reflect a proclivity for



simultaneously, and the results are compared to the base case (see Figure 1.1).

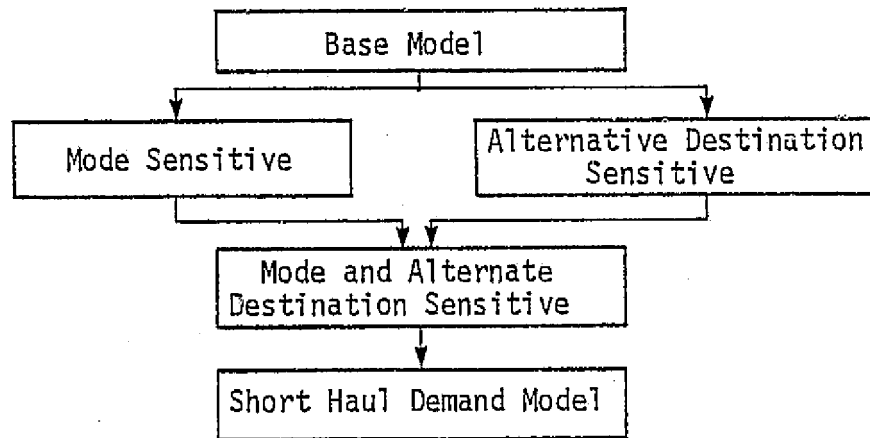


Figure 1.1 Development of the Short Haul Demand Model

The empirical support required in the model testing process is discussed in Appendix A and B. Appendix A describes the data sources while Appendix B reviews the variables, how they were constructed, and why. With the use of these data in the background, the next chapter is a discussion of the development and specification of each model followed by an evaluation of each one's statistical results. The final chapter in the report summarizes the conclusions derived from the model results and their implications for short haul air transportation demand analysis.

---

travelers to be sensitive to modal and destination differences: hence the use of the term "sensitive" to separate these models from the more traditional "split" or market share models.

## SECTION 2

## MODEL SPECIFICATION AND SEQUENTIAL RESULTS

## 2.1 Base Model

Like most intercity demand models, this short haul model begins with a gravity specification. Since the gravity approach has met notable success in past studies and no other specification has achieved equal results, it is a natural starting point. The base model is represented by the following two equations:

$$T_{ij} = b_0 G_{ij}^{b_1} I_{ija}^{b_2} \quad (2.1)$$

$$G_{ij} = M_i \cdot M_j \cdot D_{ij}^{a_1} \quad (a_1 = -1) \quad (2.2)$$

where

- $T_{ij}$  = total number of origin-destination air travelers between cities i and j
- $G_{ij}$  = travel attraction between cities i and j
- $I_{ija}$  = disutility of air travel between cities i and j, also called air impedance
- $M_i$  = effective buying income of city i
- $M_j$  = effective buying income of city j
- $D_{ij}$  = distance between cities i and j
- $b_0, b_1, b_2$  = constants determined by regression
- $a_1$  = described in section 2.1.3

## 2.1.1 Gravity Term

A city's mass or size should correspond to its population and its residents' propensity to travel. The propensity to travel refers to an individual's desire to travel and financial resources. To combine the size and propensity to travel, modelers now commonly use total income. Other more complex measures have been considered which either require substantial effort to collect and process the data or need further research. Since the purpose of this effort is to explore the influence of alternate modes and alternate destinations, the study of more sophisticated mass variables was not pursued.

The gravity term combines two masses and distance to form a single variable. Other studies have decomposed the gravity variable using structures such as:

$$M_i^{c_1} \cdot M_j^{c_2} \cdot D_{ij}^{c_3} \quad (2.3)$$

For example Alcay used  $M_i^{c_1} \cdot M_j^{c_2} \cdot \text{FARE}_{ij}^{c_3}$  where FARE replaced distance (1). However, to separate  $M_i$  and  $M_j$  requires that all  $i$  cities differ in a fundamental way from all  $j$  cities. In Alcay's model, all  $i$  cities were in the U.S. and all  $j$  cities were in Europe. In the present study, however, no fundamental difference exists to separate the sample cities into two groups, so the indices  $i$  and  $j$  have no meaning other than indicating two different cities are being considered. Thus decomposing the gravity term does not appear to be worthwhile in this study.

Since distance correlates with impedance and intermodal competition, it

is buried in the gravity term to reduce the effects of multicollinearity. Its exponent ( $\alpha = -1$ ) results from the Level of Interest effect, to be discussed shortly.

### 2.1.2 Air Impedance

The impedance term ( $I_{ija}$ ) equals the trip time plus the trip cost. This formulation was drawn from the United Technologies Research Center (UTRC) study on travel demand, which combined time and money costs into a single variable (34). Other costs can easily be added if their value in terms of time is known. For example, an access-egress variable could be incorporated if desired.

The specification for impedance used in this study is shown below.

$$I_{ija} = \text{Block Time} + \text{Waiting Time} + \text{Fare} \cdot V \quad (2.4)$$

$$V = \text{Hours/Dollars} = 1/\text{Value of Time} \quad (2.5)$$

Block time and fare can be defined quite easily. However, waiting time is more difficult to handle due to the lack of a universally accepted quantitative definition. Qualitatively the waiting time equals the time lost because a mode is scheduled, forcing a traveler to wait for the next available vehicle. Eriksen has developed a model which defines waiting time as the difference between the desired departure time and the actual departure time (22). His model appears to be the most advanced in the field presently, and is therefore incorporated into our model.

### 2.1.3 Level of Interest

Distance in the gravity term represents an hypothesized phenomenon labeled level of interest (or LOI), which assumes that a typical traveler has a decreasing desire to travel as distance increases, ceteris paribus. This effect follows from three assumptions: ceteris paribus (1) the desire to travel increases with the number of contacts (to be defined in a moment), (2) the number of trips decreases with impedance, and (3) the number of contacts increases with the number of trips.\*

The gravity term should account for the level of interest (LOI) effect alone, but since the distance variable correlates with LOI and intermodal competition an unwanted effect (intermodal competition) can enter the gravity specification unless special care is taken. In this study,  $a_1$  was found to equal -1, which is considered a reasonable value given only level of interest. Though not exact, fixing this constant should allow intermodal competition

---

\* Most short haul air passengers are business or personal travelers, which presumably indicates they have business or personal contacts in the destination city and are traveling because of those contacts. It seems reasonable to assume that if a person travels because of contacts, then his desire to visit a given destination increases with his contacts there. All else being equal, an individual should be expected to travel to that destination with the lowest impedance. This implies that the number of trips should decrease as impedance increases. Furthermore, since impedance increases with distance (by definition), the number of trips by all modes should also decrease as distance increases. Finally, it has been assumed that the number of contacts should increase with the number of trips. This appears reasonable since traveling to a destination is one way to make contacts.

Given that trips decrease with distance and contacts increase with trips, contacts should decrease with distance. Finally, combining this result with assumption (1), one concludes that the desire to travel decreases with distance (see Figure 2.1). Therefore, LOI measures the desire to travel, whereas the impedance measures the resistance to travel.

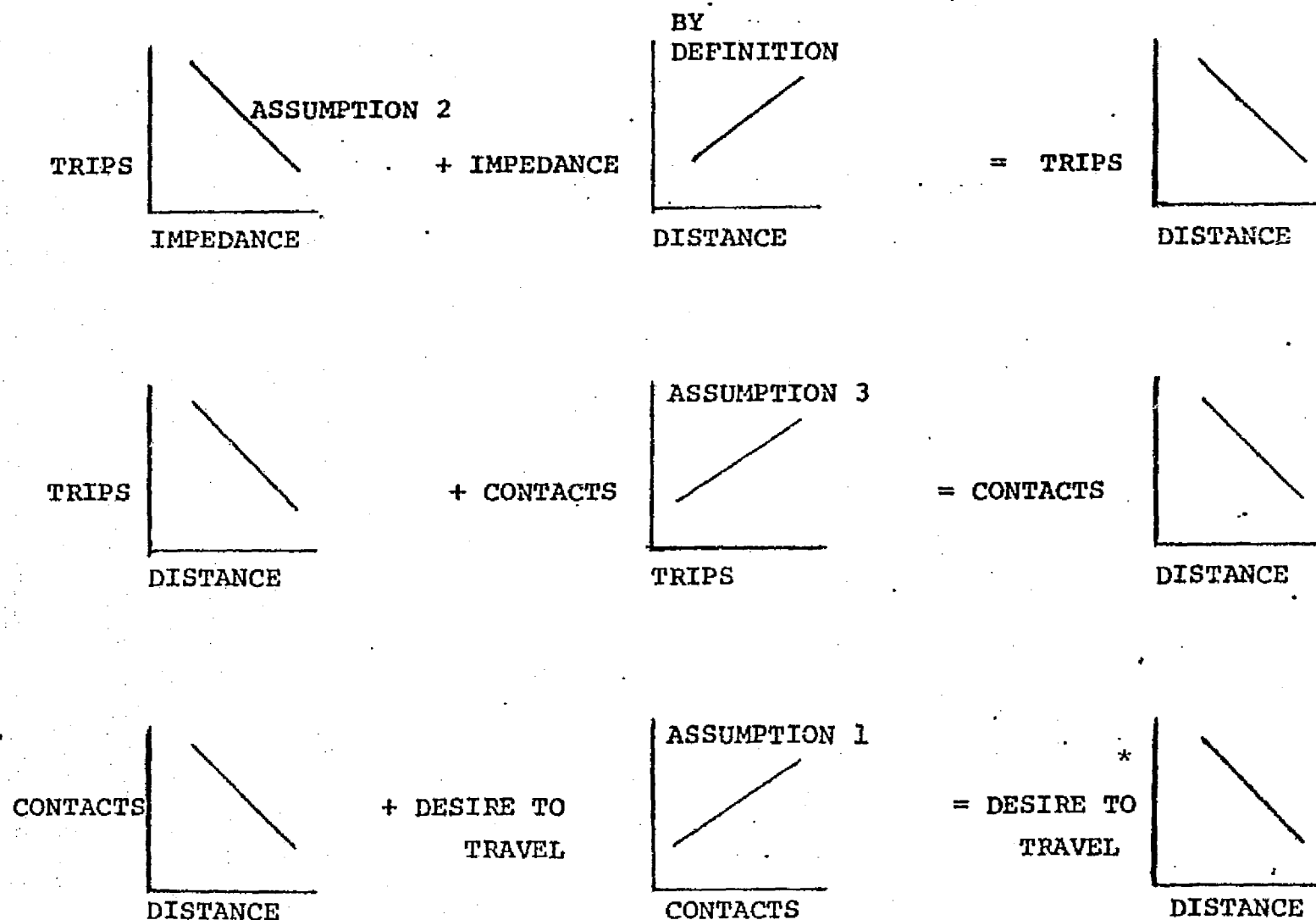


Figure 2.1. Derivation of Level of Interest Versus Distance

\* Desire to Travel is equivalent to level of interest (LOI) by definition

to be measured independently of the distance variable.

#### 2.1.4 Empirical Results of the Base Model

The base model, shown below, provided a moderate fit ( $R^2 = 0.57$ ). The coefficients had the correct signs and were all significant, and the F statistic had a value of 88 for 137 observations. The estimation of equation (2.1), or  $T_{ij} = b_0 G_{ij}^{b_1} I_{ija}^{b_2}$ , yields the following:

$$\ln T_{ij} = 10.1 + 0.32 (\ln G_{ij}) - 1.40 (\ln I_{ija}) \quad (2.6)$$

(4.9)\*                      (-6.9)\*

The base model requires a value for  $V$  (time value of a dollar) and for the exponent of distance ( $a_1$ ). To determine the optimal  $V$ , a series of values were tested in the model, and the  $V$  giving the model with the best fit was assumed to be the optimal  $V$ . As noted earlier, the exponent of distance ( $a_1$ ) is set equal to -1. As an exercise, however, an optimal  $a_1$  was determined concurrently with  $V$ , to demonstrate air traffic increasing instead of decreasing with distance, and confirm the hypothesized effect of intermodal competition. Note that LOI and intermodal competition have opposite effects on air demand relative to distance.

Numerous calibrations of the base model yielded  $V = 0$  and  $a_1 \approx 2$  (see Table C.1). The exact value of the distance exponent is not important since the objective is to determine its sign only. The value of  $V$  equal to zero

---

\* t coefficients

implies an infinite value of time, which is obviously not the case. It really shows that for the level of fares experienced in short haul markets and the type of traveler in these markets, the fare (at the current level) is insignificant. This conclusion is not so unrealistic considering who travels by air in short haul markets (100-400 miles). A one way fare in the 10-50 dollar range may be insignificant to a business traveler whose time is important and whose company pays the bill. Of course, if fares were to be raised continuously, at some level they would begin to suppress demand. In conclusion, the demand in short haul markets is asserted to be insensitive to small changes in fares at their present level.

The optimal exponent of distance is positive, which confirms the work performed earlier at M.I.T. (18). The fare in short haul markets was found to have a positive elasticity, implying that demand increases as fare increases. Actually, demand for air transportation increased with distance, not fare, but in that model, fare absorbed the effects of distance because of their strong correlation. The positive value is hypothesized to result from mode competition. Air gains a competitive edge as distance increases and therefore air traffic increases with distance in short haul markets. The author of the earlier M.I.T. work recognized this situation:

The impact of the existence of alternative modes, which are not accounted for in the model, rendered the estimation of the coefficients in the ultra short haul category model questionable. Most notable is the spurious positive correlation between fare and demand (although one might argue that the income effect is so strong here, that the coefficient should be positive). Within the range of zero to 160 miles, as the stage length decreases, air travel becomes less attractive due to the alternative of surface transportation. So in this category there is a situation where demand and fares both increase as a function of length of haul. The statistical result was a dubious price elasticity estimate of +0.9346 (see Reference 18).



## 2.2 Mode Sensitive Model

Normally one expects demand to decrease with distance, due in part to decreasing level of interest and in part to increasing impedance (travel time and cost). The results in the previous section, however, suggest that the optimal exponent of distance is positive, implying that demand increases with distance. This inference contradicts classical transportation theory, unless one also recognizes competition from other modes. Auto, for example, has a substantial time advantage over air in very short haul situations, because one need not wait for the next scheduled auto, since it is always ready. Therefore auto captures most of the very short haul market (<100 miles). As the distance increases, the speed by air compensates for its waiting time, giving the air mode a significant advantage. As its competitive advantage grows, the air mode captures a larger share of the total trips by all modes, thereby increasing its market share.

The mode sensitive model attempts to capture the effects of competing modes on the demand for air services. Two specifications are evaluated, the first simply adding a proxy to the base model, the second calculating total travel by all modes and then the share of the total carried by the air mode. Both specifications appear in detail below.

### 2.2.1 Proxy Specifications

Proxies related to the advantage of air travel relative to other modes combine with the base model to produce the proxy specification of the mode sensitive model.

$$T_{ij} = b_0 G_{ij}^{b_1} I_{ija}^{b_2} CP_{ij}^{b_3} \quad (2.7)$$

$CP_{ij}$  = Proxy for level of intermodal competition

Before defining the proxies, the term "harmonic mean" should be discussed. It does not correspond to the harmonic mean found in most statistical textbooks. Instead it derives from a United Technologies Research Center (UTRC) demand study (35). Let  $I_r$  be the impedance of the  $r^{th}$  mode and let  $\bar{I}$  be the harmonic mean of the impedances, then:

$$\frac{1}{\bar{I}_{ij}^2} = \sum_{r=1}^R \frac{1}{I_{ijr}^2} \quad (2.8)$$

United Technologies Research Center reported:\*

The harmonic mean represents the overall disutility of travel, considering all modes. It is always less than the lowest mode disutility, but is very near the lowest disutility if all other disutilities are much higher. Without the exponent, it is completely analogous to the overall electrical impedance of several impedances in parallel, an apt analogy since the traveler ("current") can choose any one mode ("impedance") to complete his journey. The exponent of 2.0 was found to improve the model correlation in early studies by keeping the harmonic mean closer to the lowest disutility (35).

The harmonic mean has several desirable properties not mentioned in the quote above. All modes are combined into a single number, the total transportation impedance, called the harmonic mean. This structure helps to avoid correlations common to models entering each mode separately and

---

\* Disutility as defined by UTRC has the same meaning as impedance.

avoids the definition of "best" or "average" travel time and fares (versus the abstract mode models). Also, a missing mode, such as rail for many of the city pairs, can also be considered. The missing mode simply does not appear in equation (2.8). Two harmonic means appear in the list of proxies. The first ( $\bar{I}_{ij}$ ) combines air, auto and rail while the second ( $H_{ij}$ ) combines only auto and rail: for  $\bar{I}_{ij}$  = Total Transportation Impedance is obtained

$$\frac{1}{\bar{I}_{ij}^2} = \frac{1}{I_{ija}^2} + \frac{1}{I_{iju}^2} + \frac{1}{I_{ijr}^2} \quad (2.9)$$

and for  $H_{ij}$  = Impedance of other modes (other than air) is obtained

$$\frac{1}{H_{ij}^2} = \frac{1}{I_{iju}^2} + \frac{1}{I_{ijr}^2} \quad (2.10)$$

where  $I_{iju}$  = auto impedance and  $I_{ijr}$  = rail impedance, respectively. The complete list of proxies is shown in Table C.2. This list is adequate although by no means exhaustive.

The correlation between distance ( $D_{ij}$ ) and intermodal competition has already been noted, and as such it is an obvious variable to be included in the model. The other three single type proxies ( $I_{iju}$ ,  $H_{ij}$ , and  $\bar{I}_{ij}$ ) describe various combinations of transportation services. Auto is the air mode's single largest competitor,  $H_{ij}$  takes into account all modes other than air, and  $\bar{I}_{ij}$  combines all modes. The Ratio Type proxies and Difference Type proxies go one step further and compare the air mode to the Simple Type proxies.

### 2.2.2 Air Share Specification

The gravity formulation predicts aggregate travel (travel by all modes) as well or better than disaggregate travel (travel by one mode, say air)(2). Thus, instead of forecasting air traffic directly, one might forecast total traffic first and then forecast the fraction of that traffic captured by air. Urban modelers use this concept regularly for their mode split models. Unfortunately, since comprehensive intercity trip data for modes other than air is unavailable, a true modal split model could not be developed here.

Consider the base model with its gravity and air impedance terms. If one assumed the gravity term represents aggregate travel (travel by all modes) and substitutes the harmonic mean of all modes (aggregate travel impedance) for air impedance, an aggregate travel model could be developed. To produce trips by air a third term is added to the model which relates to air's share of the market (see Table C.3). The final result is the air share specification:

$$T_{ij} = b_0 G_{ij}^{b_1} \bar{T}_{ij}^{b_2} f_{ij}^{b_3} \quad (2.11)$$

where  $f_{ij}$  = share of the travelers using the air mode.

The variables representing the fraction are identical to the proxies used in the proxy specifications. More complex variables might be formulated, particularly those producing probit error curves, but no attempt to that end is made here.

### 2.2.3 Proxy Specification Results

The distance and difference type proxies produced the best models, yielding the highest  $R^2$ , the highest  $t$  ratios and the highest  $F$  statistics. The remaining simple type proxies performed moderately well, but fell just short of distance and difference in all three statistical measures. The ratio type proxies matched the  $R^2$  and  $F$  of the simple proxies, but due to a correlation with  $I_{ija}$ , the  $t$  statistic of  $I_{ija}$  decreased substantially, a possible result of multicollinearity (see Table C.4).

The coefficient of the gravity term ( $G_{ij}$ ) held impressively stable throughout the calibrations, ranging between .31 and .38. The explanatory power of  $G_{ij}$ , therefore, appears insensitive to the choice of proxies used to represent intermodal competition. The coefficient of the impedance of air, however, was not stable, fluctuating from -.24 to -4.57, with the true value probably being between -1.64 and -2.19 (after adjusting for statistical abnormalities). The ratio type proxies produced low coefficients, but those models were poor statistically because of the obvious multicollinearity caused by the strong correlation between  $I_{ija}$  and the proxies, so the low coefficients should be ignored (see Table C.5).\*

The different type proxies behave in a unique manner, and therefore deserve further consideration. They correlate strongly with  $I_{ija}$ , are almost unrelated to the dependent variable, and yet these proxies produce

---

\* One can also discount the high coefficient (-4.57) produced by  $\bar{I}_{ij}$ , which also had a strong correlation with  $I_{ija}$ . The  $t$  coefficient of  $\bar{I}_{ij}$  dropped only slightly, but the significant increase in the error over the base case indicates some statistical weakness.

high quality models. The explanation lies in the partial correlations. The different type proxies have a very high partial correlation to the dependent variable when  $I_{ija}$  is in the model. In fact, each of the simple type proxies (except  $\bar{T}_{ija}$ ) demonstrates similar behavior, though not so obviously. Clearly, important interactions are occurring between air impedance ( $I_{ija}$ ) and the intermodal competition proxies.

One might ask if the different type proxies have not encountered problems similar to the ratio type proxies and  $\bar{T}_{ija}$ , which also correlate with  $I_{ija}$ . This is not the case. Note the inferior proxies correlate strongly with both  $\bar{T}_{ija}$  and trips, a clue that they explain the same travel behavior as  $I_{ija}$ . The better proxies, however, correlate with  $I_{ija}$  but not trips, yet have a strong partial correlation with trips with  $I_{ija}$  in the model, suggesting that they explain different travel behavior from that explained by  $I_{ija}$  (see Table C.6).

$I_{iju}$  and  $H_{ij}$  entered three proxies each, and every time  $I_{iju}$  produced superior results. This suggests that rail is not a serious competitor to the air mode, rail is less expensive, but in most cases has no time savings over the air mode. This is not conclusive proof that rail should be ignored, but does suggest that most intercity rail service in the U.S. is unattractive to the typical short haul air traveler. Rail service between New York, Washington, and Philadelphia appears to be quite favorable, and does attract numerous business travelers, but  $H_{ij}$  reflects that service and it still performed poorly compared to  $I_{iju}$ . Apparently rail service must increase in frequency and/or speed before a serious threat to the air mode takes place

(see Appendix B and the work performed by Systems Analysis and Research Corporation [21] and the Sloss-Kneafsey study [28]).

#### 2.2.4 Air Share Specification Results

The results appear similar to the proxy specifications. The distance variable performed the best, followed closely by the difference type variables. The ratio type variables produced low  $t$  statistics for  $\bar{T}_{ij}$  and therefore were rejected as statistically inferior. Presumably, the same interactions exist here as in the proxy specification. The ratio variables explain nothing new over the  $\bar{T}_{ij}$  term, while the other fraction variables have high partial correlations to trips and do explain demand not covered by  $\bar{T}_{ij}$  (see Table C.7).

The gravity term again remains steady, holding between .30 and .43, while the impedance term jumps from -.30 to -2.63. Excluding the ratio type variables, though, narrows the range of  $\bar{T}_{ij}$  coefficients (-2.05 to -2.63).

#### 2.2.5 Summary

Three specifications emerge as equally attractive, as shown in Table C.8. Signs on all coefficients are as expected, the  $F$  and  $t$  statistics are significant, and  $R^2$  has improved considerably over the base model. In fact, every statistical aspect of the model improves with the addition of inter-modal competition. The three specifications are so close that all three will be carried forward to the combined mode-destination model.

The coefficient of the gravity term appears very stable, while the impedance terms, as already noted, are less dependable.\* The impedance t statistics are significant.

Only auto is considered a serious competitor to short haul air travel. Efficient short haul passenger rail service is virtually non-existent in the U.S., and as such it is dropped out in the remaining analysis.

### 2.3 Destination Sensitive Models

The base model examines a pair of cities while ignoring all others, as if they did not exist. In fact, other cities do exist and their very presence affects the traffic between the original two cities. For example, envision persons traveling from city i to city j, and assume another city A lies nearby (Figure 2.2). If travelers in city i have a choice to make, do they travel to city j or city A? If A did not exist, then all traffic would go to j, but with A, some of the traffic is siphoned from j. So the alternate destination A has the effect of decreasing the number of trips made to j and may possibly increase the total number of trips generated by city i due to A's own unique attractions. These results are magnified if many alternate destinations are available.

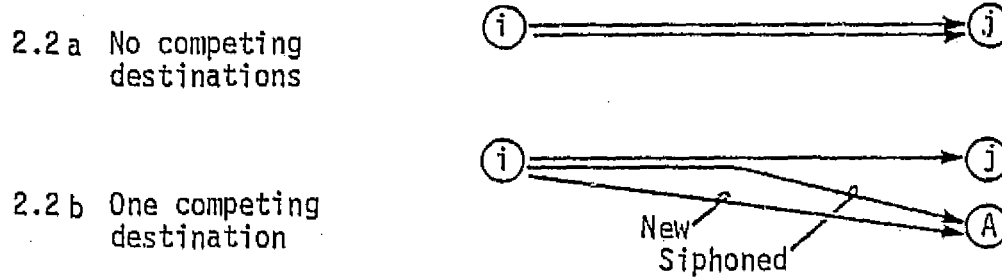
The UTRC report noted the alternative destination phenomenon when comparing city pairs in the West to those in the East (36). Apparently

---

\* The mode sensitive model assumes the exponent used in the derivation of the harmonic means equals two. To test this hypothesis, exponents equalling 1 and 1.5 were substituted into the model and tested (see Table C.9). The models were very stable, implying the exponent has little effect upon the results.



Figure 2.2 Effect of a Competing Destination on City-Pair Traffic



western cities produced more trips than expected and eastern cities produced fewer. The report concluded that the city-pairs in the sparsely populated West generated proportionally more trips because of the lack of alternatives available to travelers compared to the East. By inserting a term into their demand equation to account for alternate destinations, the differences were reduced. Since this short haul demand model samples city-pairs blanketing the U.S., it too should account for alternate destinations.

The reason for this discrepancy, which was first postulated and later validated, appears to be related to the number of travel choices available. Given a reasonably constant propensity to travel, as noted above, the trip demand between two centers having a given travel attraction (product of populations) will vary depending on the number of alternatives available. As a result of the availability of many other trip opportunities (other cities), travel between two cities in a dense region will be much less than travel between two other cities (having the same travel propensity as measured by population and distance) in a sparsely settled region.

Reference 22(3)

The alternate destination concept is relatively unexplored. So three different specifications are proposed and tested here. The first two envision each city as a source of travelers which radiate out in all directions to all destinations, and the third simply adds a proxy. The first source specification, source-one, assumes the number of travelers generated is a function of the size of the source city and the attraction of

all potential destinations. The second source specification, source-two, assumes total generation depends upon the size of the source only.

The source specifications operate conceptually in two steps. The western cities produced more trips than expected and eastern cities produced fewer. The report concluded that the city-pairs in the sparsely populated West generated proportionally more trips because of the lack of alternatives available to travelers compared to the East. By inserting a term into their demand equation to account for alternate destinations, the differences were reduced. As the UTRC group observed:

The reason for this discrepancy, which was first postulated and later validated, appears to be related to the number of travel choices available. Given a reasonably constant propensity to travel, as noted above, the trip demand between two centers having a given travel attraction (product of populations) will vary depending on the number of alternatives available. As a result of the availability of many other trip opportunities (other cities), travel between two cities in a dense region will be much less than travel between two other cities (having the same travel propensity as measured by population and distance) in a sparsely settled region (36).

Since this short haul demand model samples city-pairs blanketing the U.S., it too should account for alternate destinations.

### 2.3.1 Model Specifications

The alternate destination concept is relatively unexplored, so three different specifications are proposed and tested here. The first two envision each city as a source of travelers which radiate out in all directions to all destinations, and the third simply adds a proxy. The first source specification, source-one, assumes the number of travelers generated is

a function of the size of the source city and the attraction of all potential destinations. The second source specification, source-two, assumes total generation depends upon the size of the source only. The three specifications are depicted below.

The source specifications operate conceptually in two steps. A source term generates total travelers departing by the air mode to all destinations, and a second term calculates the fraction of those travelers flying to the other source city. The source-one specification assumes generation corresponds to city size and the attraction of all potential destinations, or  $M \times S$  where  $S$  equals the strength of attraction of all destinations for a city.  $S$  is defined below. The source-two specification limits generation to a city's mass or size, and therefore equals  $M$ . The fraction can be derived directly from the generation terms.\*

The proxy specification is very simple. Three proxies were derived using the destination attraction variable  $S$ , and each tested in the base model. This approach is not as elegant as the source method, but is an obvious approach that should be covered for completeness.

$$\text{Source-One: } T_{ij} = b_0^{GH_{ij1}} b_1^{I_{ija}} b_2^{F_{ij1}} b_3 \quad (2.12)$$

$$\text{Source-Two: } T_{ij} = b_0^{GH_{ij2}} b_1^{I_{ija}} b_2^{F_{ij2}} b_3 \quad (2.13)$$

$$\text{Proxy: } T_{ij} = b_0^{G_{ij}} b_1^{I_{ija}} b_2^{P_{ij}} b_3 \quad (2.14)$$

---

\* For a detailed derivation of the generation and fraction terms, see Blumer's thesis (5), Appendix A.2.

where

$$GH_{ij1} = [(M_i \times S_i)^{\theta_1} + (M_j \times S_j)^{\theta_1}], \text{ (Source-One)}$$

air travelers generated from cities i and j

$$GH_{ij2} = (M_i^{\theta_2} + M_j^{\theta_2}), \text{ (Source-Two)}$$

air travelers generated from cities i and j

$$S_i = \text{total attraction of alternate destinations for travelers from city i, } S_i = \sum_{\text{all destinations}} M_j / D_{ji}$$

$$S_j = \text{total attraction of alternate destinations for travelers from city j, } S_j = \sum_{\text{all destinations}} M_i / D_{ij}$$

$\theta$  = exponent on generation terms, as generation may not be directly proportional to the source strength

$$F_{ij1} = \text{fraction of total travelers moving between cities i and j (Source-One specification)}$$

$$F_{ij2} = \text{fraction of total travelers moving between cities i and j (Source-Two specification)}$$

$$G_{ij} = M_i \cdot M_j \cdot D_{ij}^{-1}$$

$P_{ij}$  = proxy representing alternate destinations. The three proxies used are  $S_i \cdot S_j$ ,  $S_i + S_j$ , and  $W_i S_i + W_j S_j$ . The latter is a weighted average [see Blumer (5)].

### 2.3.2 Statistical Results

Overall the results were mixed (Table C.10). The Source-Two specification increased  $R^2$ , but reduced the  $I_{ija}$  statistics nearly in half. The proxy specification affected  $I_{ija}$  in the same way and produced even less improvement in  $R^2$ . The worst case was Source-One, not only did the

t coefficient of  $I_{ija}$  drop, but  $R^2$  made no improvement over the base case.

The exponents within each case had little or no effect. The exponent ( $\theta$ ) produced almost identical R's over a range of values. Meanwhile the generation coefficient decreased in magnitude as  $\theta$  increased, but its t coefficient held steady. The fraction term and  $I_{ija}$  appeared unaffected. The exponent  $\theta_2$  behaved in an identical manner to  $\theta_1$ . The three proxies produced almost identical results, making a choice among the three arbitrary.

These results depend strongly upon the calculation of S. Up to this point the exponent of distance in the S calculation was assumed to equal one. However, this assumption should be tested. So exponents ranging from 0 to 2 were tested, and the results showed the model was insensitive to this value (Table C.11). Only the S coefficient changed, and the changes were very small. Thus, an exponent of one will continue to be used.

The source-two and proxy specifications will be carried to the mode-destination phase of investigation in the next section. The source-two specification provided the best results and was therefore the natural choice for the destination sensitive model. The proxy specification provided moderate results. However, its simplicity is quite appealing.

The destination sensitive model shows little improvement over the base model.  $R^2$  increased slightly, F was as high or higher, the t coefficient of gravity (in the proxy specification) increased, and the t coefficient of  $I_{ija}$  decreased. With the tradeoff between  $R^2$  and the t coefficient of  $I_{ija}$ , the results of the base model and mode sensitive model appear quite similar.

## 2.4 Mode-Destination Sensitive Model

### 2.4.1 Model Specification

Three specifications emerged from the mode sensitive model and two more from the destination sensitive model. By combining the simpler specifications, six mode-destination sensitive models can be derived. These are the final six models investigated in this report. Each is briefly described below.

$$T_{ij} = b_0 G_{ij}^{b_1} I_{ija}^{b_2} D_{ij}^{b_3} S_{ij}^{b_4} \quad (2.15)$$

$$T_{ij} = b_0 G_{ij}^{b_1} I_{ija}^{b_2} DA_{ij}^{b_3} S_{ij}^{b_4} \quad (2.16)$$

These specifications calculate air travelers between i and j directly. The mode and destination variables are proxies.

$$T_{ij} = b_0 G_{ij}^{b_1} T_{ij}^{b_2} D_{ij}^{b_3} S_{ij}^{b_4} \quad (2.17)$$

The gravity term represents i to j travelers by all modes and  $D_{ij}$  is assumed to correlate with the fraction of those travelers moving by air.

$$T_{ij} = b_0 GH_{ij2}^{b_1} I_{ija}^{b_2} D_{ija}^{b_3} F_{ij2}^{b_4} \quad (2.18)$$

$$T_{ij} = b_0 GH_{ij2}^{b_1} I_{ija}^{b_2} DA_{ija}^{b_3} F_{ij2}^{b_4} \quad (2.19)$$

Now a source term appears instead of a gravity term. It represents all

air travelers generated at  $i$  and  $j$ , and  $F_{ij}$  represents the fraction traveling between  $i$  and  $j$ . The mode variable is a proxy.

$$T_{ij} = b_0 GH_{ij2}^{b_1} \bar{T}_{ij}^{b_2} D_{ij}^{b_3} F_{ij2}^{b_4} \quad (2.20)$$

All travelers by all modes are represented by  $GH_{ij2}$ . The variable  $F_{ij2}$  embodies the fraction moving between  $i$  and  $j$ , and  $D_{ij}$  correlates with the fraction of the remaining travelers using air.

#### 2.4.2 Results of the Combined Model

None of the six specifications is clearly superior, making the choice of a model almost arbitrary. The three source specifications have higher  $R^2$ 's but the three gravity specifications have higher  $t$  coefficients for  $D_{ij}$  and  $DA_{ij}$ . The choice is further complicated because the coefficients of  $I_{ija}$  and  $D_{ij}$  vary widely, even though all  $t$  ratios are high (see Table C.12).

Rather than force a decision with the current set of data, all six models (Table C.13) were subjected to detailed statistical tests to insure that least square regression assumptions were met and that the data base was internally consistent (section 2.4.3).

As a group, the source specifications had higher  $R^2$  and  $F$  values, while the gravity specifications yielded higher coefficient and  $t$  values for  $I_{ija}$ ,  $\bar{T}_{ij}$ , and  $DA_{ij}$  (see Table C.13). Furthermore,  $DA$  produced higher  $t$  statistics for  $I_{ija}$  compared to  $D_{ij}$ , but in the gravity specification group had a lower  $F$  value compared to  $D_{ij}$ . Finally  $\bar{T}_{ij}$  produced a higher coefficient, and a better  $t$  statistic (compared to  $I_{ija}$ ), but in the source group had a lower  $F$  statistic.

Comparing the mode-destination model to the mode sensitive and the destination sensitive models proved interesting. Investigating the mode sensitive model first, one notices that the gravity coefficient increases from .40 to .60 while the  $t$  statistic remains constant. Meanwhile the  $I_{ija}$  coefficient and  $t$  statistic decrease slightly. The  $D_{ij}$  and  $DA_{ij}$  terms also drop slightly. At the same time, the explained variation increases from .81 to .84 and  $f$  decreases slightly. The tradeoff appears between  $R^2$  and the  $t$  coefficients, a small improvement in  $R^2$  for a small reduction in the  $t$  ratios.

The mode-destination model is a large improvement over the destination sensitive model. Not only do  $R^2$  and  $F$  increase significantly, but the  $t$  coefficient of  $I_{ija}$  increases from 4 to about 10 and  $t$  coefficient of  $S$  from 4 to 5. The combined model therefore reflects a substantial improvement over the destination sensitive model, but only a marginal improvement over the mode sensitive model.

#### 2.4.3 Statistical Verification for Linearity, Heteroscedasticity, Normality, and Pooled Data

The least squares regression technique used to calibrate the model specifications relies on several assumptions. The error or residuals should have a constant variance; if not, then heteroscedasticity exists and the least squares regression technique may not produce the desired properties. The estimated coefficients will be unbiased and consistent, but will not be minimum variance unbiased estimators. Therefore the estimated variances of



the model coefficients will be biased, and if used, the statistical tests and confidence intervals will be incorrect. In addition, the model is assumed to be linear (in the log transform). If it is not, then the fit is poorer than it should be and forecasts may be unreliable. Finally, the error or residuals are assumed to be distributed normally with a mean of zero. Again, if this condition does not hold, most of the common statistical tests are no longer appropriate.

A scatter plot provides a simple check on both linearity and constant error variance. The residuals are plotted against the independent variables and against the estimated values of the dependent variable. A linear distribution would lie evenly about the horizontal axis, while a non-linear distribution forms a curved shape. The variance is simply the spread of the sample points about the horizontal axis. If the spread is even, then the variance is constant, but if the spread is uneven a condition called heteroscedasticity exists. Scrutinizing a scatter plot provides only a rough feeling for linearity and constant error variance. More exact tests do exist; for example, F tests work for both conditions (23). Scatter plots were produced for all six mode-destination specifications and did not suggest any serious statistical problems [see Blumer (5), Appendix A.3].

A final test was performed on the data itself. Cross sectional data were collected for three consecutive years, and the three years were pooled. If the model parameters are constant over time, then pooling may be acceptable. To test this hypothesis, each year was regressed separately and the coefficients compared for stability, as shown in Table C.14. A Chow test can be performed if necessary. The coefficients for 1972 and 1973 matched closely, but those for 1974 appeared quite different. The question

arises as to why 1974 was so different: is the model misspecified or are the data biased by extraordinary events that occurred during 1974? The latter possibility appears likely since 1974 was the year of the fuel crisis. During this period the airlines cut back flights to reduce fuel consumption, the public traveled less, and those who did travel shifted out of their cars to public modes for fear of unavailability of fuel.

To test the fuel crisis hypothesis, the model coefficients were checked for behavior consistent with 1974's events. The short haul model assumed that all persons had the same propensity to travel, regardless of the size of the origin [see Blumer (5), Appendix A.1]. During the fuel crisis, however, people traveled less. Since a densely populated area has more local amenities than a less dense area, the dense area's residents would have less need to travel and therefore less propensity to travel. In other words, although the number of potential travelers increases with gravitational attraction, a situation like the fuel crisis would cause the probability of an individual traveling to decrease with city size. It is likely that this was a significant factor which caused the gravity coefficient for 1974 to decrease.

The impedance coefficient, however, increased. Most auto travelers use air as an alternate mode. The auto usually has distinct advantages both in trip time and convenience for short haul travel. However, during 1974 auto lost some of its convenience appeal. Travelers became fearful of finding no fuel when they ran low, and at best would have to sit in long lines waiting for a turn at the pump. As the auto lost its non-time related advantage, travelers compared auto and air strictly on the basis of time. In other words travelers became more sensitive to trip time (impedance) and

therefore the 1974 impedance coefficient in the equation for air travel demand was larger than its corresponding value in 1972 and 1973.

Distance correlated strongly with the competitive advantage of air relative to auto. The fuel crisis did not change trip time by air, but it did change that for the auto slightly. Not only did the speed limits drop from 70 to 55 but travelers had long waits at the fuel pumps. Thus, the relative advantage for air travel increased which meant the coefficient of distance could be expected to increase. In fact, the coefficient changed very little. However, it should be noted that the coefficient equals the elasticity of distance which refers to the percent change of the relative advantage for the air mode. The percent change may increase very little while the absolute change may be significant.

The fourth variable in the short haul model measures the attraction of alternate destinations. Its coefficient decreased for the 1974 calibration and in fact became insignificant. The term refers to a traveler's choice, the option to travel to any destination. The fuel crisis curtailed this behavior because people traveled only when they had to and then to a specific destination. The option to choose destinations diminished, as did the coefficient.

Given the fuel crisis which impacted so heavily upon the U.S. economy and its population, it should be no surprise that 1974 is unique compared to 1972 and 1973. In addition, the movement of the coefficients appears consistent with expected behavior. Since 1974 appears to be a unique year, drawn from a different population than 1972 and 1973, the six mode-destination specifications were calibrated on the combined 1972-1973 data set (see Table C.15). The coefficients behave similarly as in the three

year sample. The gravity and alternate destination terms increased while the impedance and competition terms decreased slightly. In addition the F values decreased very little, even though 1/3 of the sample points are eliminated and the explained variation ( $R^2$ ) increased from .86 to .90. Statistically the model calibrated on two years of data appears slightly better than that calibrated on three years of data.

In summary, our short haul model specification appears to be statistically valid. A pooled sample of observations from 1972 and 1973 produced statistical results which adhered to the properties of econometric verification. While additional observations over longer periods of time would be desirable, the initial statistical results of our model specifications appear to be structurally sound.

### SECTION 3 CONCLUSIONS

#### Model for Prediction and Analysis

A structurally sound demand model should have the capability of predicting future demand as well as adhering to the explanatory powers of historical data. In regression analysis the predictive power is often measured by  $R^2$ , given that no serious statistical problems exist. Of the six specifications, the fourth model had the highest  $R^2$  (.90), with no apparent statistical problems. The short haul demand model with specification four is shown below. Since the coefficients varied somewhat between specifications and all t statistics were high, selecting any one set of coefficient values would be arbitrary. The true value of a coefficient is far more likely to lie within the range than to be around a single value. The range is narrow enough to be useful in many applications.

The best short haul model for purposes of prediction was the following:

$$\ln T_{ij} = .9171 + .9497 \ln H_{ij2} - .9510 \ln I_{ija} + 1.3195 \ln D_{ij} \\ + .4735 F_{ij2}$$

$$\sigma = .4984 \quad R^2 = .8954 \quad F = 188.35$$

For forecasting, the short haul demand model requires the anticipated city sizes, level of air service, intercity distance (a constant), and attraction of surrounding cities (those within 300 miles). The size or economic strength of cities is forecast by economic models and is usually available locally or through government sources. Air service is not

typically forecast, except by the Civil Aeronautics Board. It should, however, be noted that frequency of service increases far more slowly than demand because aircraft have been increasing in size. Thus, one might project service as an increase by a certain percent and perform a sensitivity analysis by utilizing a range of growth factors.

The model can analyze the effects of economic changes in cities, regions and in air service. It cannot, however, analyze the effects of a new mode or drastic changes in an existing mode such as rail. The only other mode the model seriously considers is auto, and even major changes in that mode cannot be analyzed. The reason stems from the existence of an excellent highway network in the U.S. Auto service is so uniformly good that it can be represented simply by distance. This lack of variance precluded the model from developing an ability to compare modes. However, if one wished to possess such an ability and did not require demand forecasting excellence, it could be achieved. Mode split and abstract mode models were developed for this purpose. Of course, acquiring the data for their calibration may present some difficulties.

The series of calibrations and tests performed for this report produced several conclusions. Auto appears to be the only serious competitor to air in the short haul markets. No conclusive test was performed, but the theoretical development and the statistical results support this hypothesis. This conclusion is also supported by the inference that the typical short haul air traveler is insensitive to fare. These travelers tend to be business oriented and are primarily concerned with time. In fact, air service, or trip time (travel time plus waiting time), is one of the more important determinants of air travel demand. A change in air service can have a

dramatic effect upon demand since air service is elastic or very nearly so (elasticity estimates ranged from  $-.9304$  to  $-1.46$ ). Finally, the existing simple correlations show distance to underlie several of the short haul explanatory phenomena. Since distance correlates with intermodal competition, level of service, fare, and travel time, special care must be taken to separate the effects of distance -- a feature which may be responsible for producing the vast array of fare elasticities that researchers have produced in the past.

### Implication of Model Results

Air travelers's strong sensitivity toward trip time in short haul markets affects government regulation, airframe manufacturers, and airlines. Regulators have long debated the appropriate fare for short haul markets and whether these markets need be cross-subsidized by the profitable long haul markets. The argument for cross-subsidization is that a fare increase would severely suppress air traffic due to intermodal competition. This report suggests the opposite, that short haul air service competes on a time basis and the fare at its current level has little impact upon demand. Therefore, cross-subsidization appears unnecessary so that the short haul markets might carry their own load.

A good part of government sponsored research (NASA) is aimed at more efficient aircraft, using less fuel and lowering operating costs. These goals are worthy from the viewpoint of fuel conservation and more efficient use of resources. This report concludes that the short haul traveler is concerned about trip time only. To stimulate the demand the trip time must

be reduced, which means faster aircraft, higher frequencies, or faster processing at each end of the flight. Since faster aircraft usually require more fuel, that avenue may be unlikely. Higher frequencies are possible, using small profitable aircraft (for example, 20-70 passenger aircraft, depending on the market). Processing efficiencies at the origin and destination points may turn out to be the most cost-effective way to proceed in order to implement the inferences of our short haul air transportation demand mode.

### Future Research

Several areas remain to be explored in order to refine and improve the short haul demand model. The access-egress time was not included here, but given the apparent importance of trip time, one quickly concludes that access-egress time should be incorporated into the model. Calculating these times may prove to be difficult, although UTRC (33) claims to use a method that requires only moderate effort. The mass variable also needs further research (Appendix B notes some possible directions). In particular, using the income distribution weighted by the propensity to fly may improve the model. The business nature of the short haul air market suggests a business indicator such as value added. Either way, cities with special attractions such as Las Vegas or Washington, D.C. need to be handled more carefully. (For example, a refined measure of hotel rooms may be useful.)

Another area to explore is a second model representing service. Obviously, the number of flights depends upon the demand and the demand in turn depends upon the number of flights. This two-way causality suggests



building a second model representing air service to be solved simultaneously with the first model. This feature would also eliminate the requirement to forecast air service because the variable would become endogenous.

Finally, the data base needs to be extended. This study started with 49 city-pairs observed over a three year period and ended with a few less city-pairs over two years. The degrees of freedom are sufficiently high, but the model may be biased toward conditions existing during 1972 and 1973. If the data base included a longer time-series (for instance, ten years), then more compelling evidence could be brought to bear with respect to the forecasts and the analysis of the underlying behavior of demand travel in short haul air transportation markets.

PART II

An Analysis of Short Haul Air Passenger Demand

by

William M. Swan

## SECTION 1

### INTRODUCTION

Part 1 of this report has demonstrated that short haul air travel demand models could, from a statistical point of view, successfully incorporate the effect of alternative destination and modal splits, thus providing an improvement to gravity models for intercity demand markets. The objective of the work reported in Part 2 is to extend the work discussed in Part 1 by using a more causal formula than the log-linear Cobb-Douglas model.

As suggested in Part 1, such a model can be used as an analysis tool to forecast demand under altered conditions of cost, competitive mode performance, and demographic shifts. A model which incorporates cause and effect logic, i.e., a behavioral model, appears as the most appropriate tool for such policy analysis under conditions which may be different from those used initially to calibrate the model.

The purpose of this part is to suggest that it is both desirable and possible to improve on the traditional Cobb-Douglas formulation used to predict the demand and modal split for intercity trips. We intend to review the traditional mathematical form with special attention to areas where the algebra causes trouble. Then we will present a new model developed from a different concept of behavior. The new formulation will overcome the objections raised against the traditional models. Finally, we will describe a calibration of the new formula and suggest some interesting further developments.

The new formula is based on modeled rational decision-making behavior

on the part of consumers. The concepts result in a somewhat complicated non-linear formula which includes a completely disaggregated treatment of consumer demand. Calibration of such a model did not prove computationally difficult.

## SECTION 2

## TRADITIONAL MODELS

By "traditional demand models", we mean models which predict the total travel between two cities from some form of the gravity formula and the distribution of this traffic among modes using ratios of modal impedances.

The gravity model for trip generation takes the form:

$$D_{ij} \sim P_i^x P_j^x / d_{ij}^y \quad (1)$$

Here  $D_{ij}$  is the total travel between cities  $i$  and  $j$ . The numerator on the right is the product of the cities' populations. Total city incomes, disposable incomes, business activity, or other indices of wealth or activity can be used in lieu of raw populations, but the concept is the same. Traffic is generated in proportion to the originating population, and the destination population is used as an index of the number of reasons to go to that city. Exponents ( $x$  above) are used to improve the statistical fit.

The denominator in the gravity model is either intercity distance or some other index of the difficulty of getting from  $i$  to  $j$ . Greater distance (or cost) discourages travel on two counts. First it raises costs in relation to other cities, influencing destination choice. And second, higher prices discourage trips in general. In keeping with economic modeling traditions, this cost has an elasticity,  $y$ .

In its favor, the classical gravity model combines the three most relevant variables so that all local behavior has the right sign: if either city doubles in population, demand grows. Where distance is greater,

demand is diminished. Another advantage of the gravity model is that it can be calibrated (in its logarithmic form) using linear regressions such as may be done on the oldest and simplest computers. The simplicity of expression and calibration may account for this model's historical dominance. Unfortunately, there are one major and one minor objection that must be raised against the specific algebraic form.

The major problem with the gravity model involves the implications about per capita travel from a city to all destinations. We represent this per capita total travel as:

$$\frac{D_i}{P_i} = \frac{\sum_j D_{ij}}{P_i} \sim P_i^{(x-1)} \times \sum_j (P_j^x / d_{ij}^y) \quad (2)$$

For the per capita travel from  $i$  to be nearly independent of that city's population ( $P_i$ ),  $x$  must be near one.\* But if  $x$  equals one, the per capita travel is directly proportional to the national population ( $\sum_j P_j$ ). Conversely, for per capita travel to be nearly independent of the national population,  $x^{**}$  must be zero. This makes per capita travel inversely proportional to city size ( $P_i$ ) and intercity demand independent of population. It would be a coincidence if the two tendencies (for  $x$  to be near 1 and near zero) could be satisfied by some middle value such as  $x = 0.5$ .

---

\* Per capita travel may be correlated with city size, but the model ought to be able to handle the independent case since it is a reasonable first order assumption.

\*\*

Again, independence is merely a desirable possibility, because it is a reasonable first order assumption.

Unfortunately, it is just such a coincidence which any statistical calibration of the gravity model assumes, and hence obtains.

The second objection to the gravity model is minor and we mention it only for completeness. Mathematically, there is a singularity at the origin ( $d_{ij} = 0$ ). This is not of practical interest except that we must always beware that statistical calibrations are not dominated by observations at the shortest distances. The singularity will highly leverage these data points and a poor fit at longer distances may occur.\*

Difficulties with the traditional modal split term are more subtle, but they are more likely to mislead policy decisions. The traditional modal split defines a conductance  $K_m$  for each mode,  $m$ , as some function of the mode's cost  $C_m$ , travel time  $T_m$ , and other characteristics,  $A_m$ . Usually,

$$(i) \quad K_m = A_m C_m^\gamma T_m^\sigma \quad (3)$$

or

$$(ii) \quad K_m = A_m e^{(\gamma C_m + \sigma T_m)} \quad (4)$$

Modal splits are then made for  $n$  according to the formula:

$$F_n = \frac{K_n}{\sum_m K_m} \quad (5)$$

where  $F_n$  represents the fraction of traffic allocated to  $n$ , and the term

\* A requirement that per capita trips to all destinations decrease with distance under conditions of uniform population density implies that  $y > x$  in formula (1). This is not, however, a criticism of the algebraic form, but merely a constraint occasionally neglected in calibration.

$\sum_m K$  is the sum of the conductances of all modes.  $\gamma$  and  $\sigma$  are negative and the (implicit) value of time for consumers in each mode may be deduced from the ratios of price and time elasticities to be  $(\sigma/\gamma) \times C_m/T_m$ .

The best known objection to equation (3) or (4) is referred to as the "red bus, blue bus" problem. Consider a case with three modes: fast, medium, and slow. Imagine the slow mode is provided by a service of buses painted red. Now add to the system a new mode, buses painted blue but otherwise operated exactly the same as the red bus mode. What does this do to the total bus market share? It nearly doubles it.\* This is an unavoidable consequence of the algebra. It is also a significant calibration problem when nearly comparable rail and bus service exist in some but not all markets. Traffic on the slow mode or modes should not double and halve depending on the modal distinction.

A second objection is nearly the same as the first, but in a different disguise. Consider the three mode case above. Now imagine a fourth mode yet slower and cheaper than bus travel. According to the formula, this new mode will capture an equal fraction of the slow, medium, and fast mode customers. Intuitively we prefer modes to have small competitive effects on the demand for extremely different qualities of service. Improvements in performance of one mode should largely affect the demand only of the adjacent modes. Policy tools unable to demonstrate such behavior would seem questionable.

While these objections have been illustrated with extreme cases, the lack of appropriate algebraic behavior throws into question the use of such models

---

\* For 10% red bus travel at start, red and blue get 18%.



except as a means of interpolating in the vicinity of observed data points. Researchers have long recognized these problems, especially with respect to intraurban trip modeling. One of the most fruitful approaches has been to disaggregate the modal split problem into as many classes of consumers as possible, and then perform a traditional modal split for each class. The new formula we propose takes this process to its logical conclusion.

## SECTION 3

## A NEW FORMULA

The proposed new demand formula will model four multiplicative effects to predict the demand for each mode in a city-pair market. Total travel generation<sup>\*</sup> to all points from the origin will be predicted in the traditional way from population, income, business activities, or some index reflecting the numbers of people involved and their propensity to travel. The distribution of this demand among potential destinations will be made considering the effects of intervening destination choices. The fraction of demand employing each mode will be predicted from a continuous disaggregated distribution of consumer classes. And finally, consumption of each mode will be adjusted according to a nearly traditional form of cost elasticity. The section below presents these components first in qualitative and then in quantitative form.

## 3.1 Travel Generation

The statement of per capita travel generation from a city implicit in the gravity model is suitable. Total travel should be proportional to a city's population, income, or some other index of the size and activity of the population. It may be that certain cities contain industries which rely heavily on travel, or that people in medium-sized towns travel less than

---

<sup>\*</sup> Strictly speaking, we predict the need for trips, some of which is not satisfied due to high prices.

those in larger or smaller ones. (It may even eventually be possible to describe such variations by the selection of appropriate indices to represent city size.) However, population should certainly be considered. To the extent that income combines the effects of population and wealth in generating travel, it would seem to be the first improvement over population.

### 3.2 Alternative Destinations

The total demand for travel must be distributed among the available destinations. As a model of general behavior, it is hypothesized that each person in the country has an equal probability of being able to serve a trip's purpose, i.e. he may be a suitable destination. At the same time it is assumed that people will travel to the closest person who satisfies the need. For example, assume that (on the average) there is one kidney specialist per million of population. There is a good chance of traveling across town to see him for residents of New York City, but one will probably travel across the state to see him if one lives in Nebraska. The probability that a trip generated at city  $i$  goes to city  $j$  is the probability that no one closer to  $i$  than city  $j$  (including other people at  $i$ ) can satisfy the trip purpose multiplied by the probability that some one in city  $j$  does satisfy the trip purpose.

This conceptual model can be stated simply in mathematical terms. For an individual living at the center of a region of uniform population density, the resulting distribution of trips against distance produces a Raleigh

distribution such as in Figure 3.1<sup>\*</sup>. This model has the interesting characteristic that there is a tendency for trips to shorten as population grows since there are more nearby people to satisfy trip purposes. This behavior has been observed for truck shipments.<sup>\*\*</sup> (There is also a probability of not satisfying a trip purpose with any member of the country's population. This probability is, however, quite small.)

The fundamental variable in this formulation is the probability of a randomly selected person satisfying a trip's purpose. The number of kidney specialists per million of population will change with time as the degree of specialization in the economy changes. On the other hand, per capita trips to visit relatives or specific places such as the Grand Canyon, for example, will not increase much as population grows. Further, some population clusters, such as those living near the Grand Canyon, appear artificially attractive as trip destinations and others, such as older industrial towns, appear less attractive. This suggests a city-by-city adjustment factor or the use of some attraction index other than population.

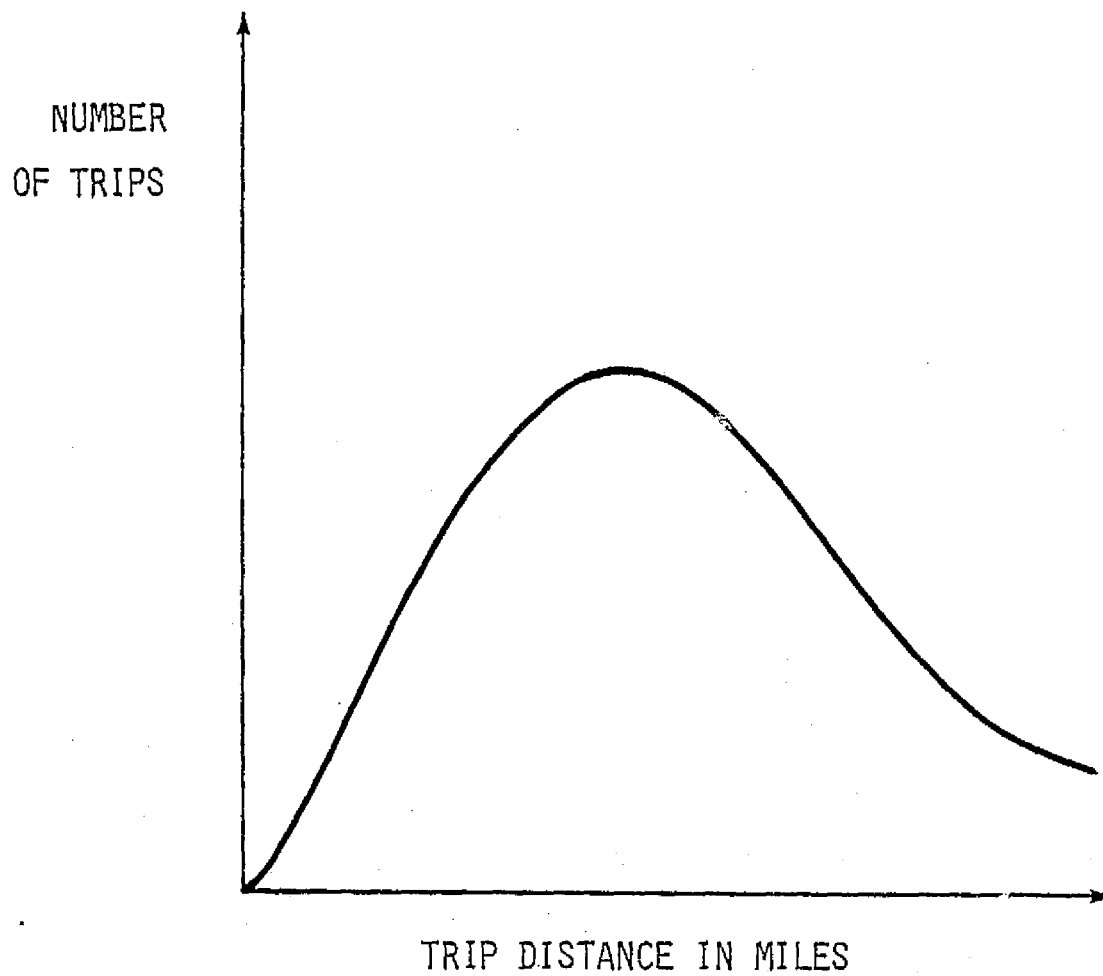
For the initial formulation, the model will be defined without time or city adjustment factors. Each person is equally likely to be a suitable destination, and the closest suitable destination is always chosen. The development of this probabilistic concept allocates travel from city  $i$  to city  $j$ . If there is a large population between  $i$  and  $j$ ,  $j$  is a less likely destination than if few people live between  $i$  and  $j$ .

---

<sup>\*</sup>This result is used for Urban Service Systems; see Larson and Odoni,(1).

<sup>\*\*</sup>Morton, (2) .

FIGURE 3.1



DISTRIBUTION OF TRIPS VS. DISTANCE FOR CONSTANT  
POPULATION DENSITY, INTERVENING OPPORTUNITIES  
FORMULATION.

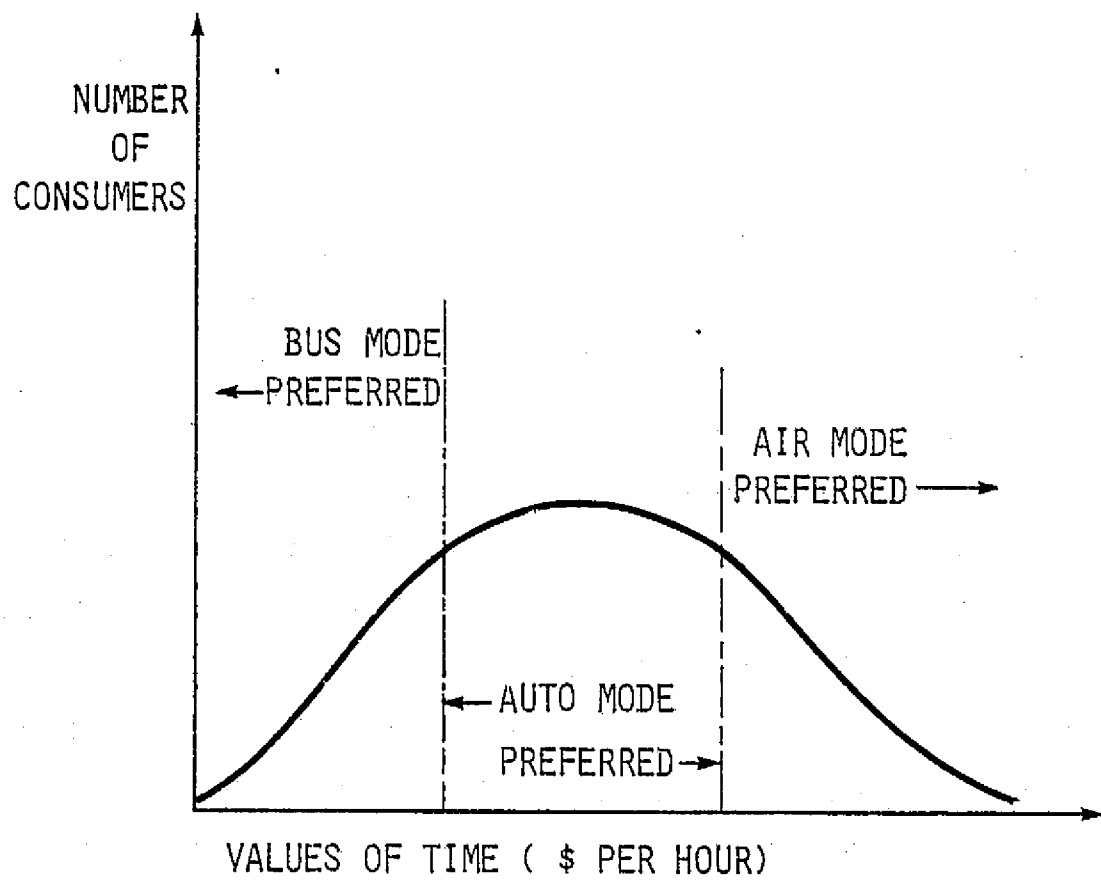
### 3.3 Modal Split

One difficulty with the traditional modal split formulation is that it implies a single value of time for all consumers. The logic which would produce such a formulation is that all people ranked the modes' price/service packages the same way, but some of them get muddled and chose the second or third best packages.

Here it is proposed that modal ranking be probabilistically distributed. That is, that people's taste be considered not all the same. Since the dominant factor is assumed to be time, it is hypothesized that there is a distribution of values of time among the population, and that every person takes the mode which is cheapest for him, including the value of his time. This modal split then becomes the process of allocating a distribution of values of time into regions preferring each mode. Figure 3.2 shows a Gaussian distribution of values of time with watersheds separating the air travelers at the high values from car travelers and car travelers from bus travelers on the low side. The car mode is assumed available to all. (In essence this is a completely disaggregated approach as each point on the value of time axis is a "cell" of people disaggregated from the rest of the market.)

Using value of time as the one dimension of people's preferences is a first approach. The concept could be more generally stated as distribution of tastes in  $n$ -dimensions. One dimension would be value of time, another would be value of having a car at the other end of the trip, and so on. These taste characteristics may be partly correlated and could perhaps be predicted from observed population characteristics such as income and profession. This surface in  $n$ -space would then be divided by cutting

FIGURE 3.2



CONCEPTUAL REPRESENTATION OF MODAL SPLIT BY VALUE OF TIME

hyperplanes corresponding to the watersheds between modes. For multiple dimensions of taste, all modes can compete with all other modes. Unlike the traditional modal split model, the degree of competition would not be fixed. Influences may be strong between one pair of modes and weak between others. For a first approach, however, only the one-dimensional case is considered where modes compete only with the modes of adjacent time/cost performance.

### 3.4 Price/Consumption

The final step of the demand model is to alter the modal consumption to reflect the total trip cost. The concept of expressing the total trip cost, including the cost of time (and other trip attributes), was developed above. Ideally this expression would be stated for each class of consumers and the appropriate price elasticities would be applied. For an initial formulation, however, it is proposed to determine the total trip cost for each mode using some representative value of time for consumers of that mode. The justification for the simplification is that a large part of price or time elasticity is already inherent in the modal split term, and only some remaining travel elasticity has to be considered with the total cost term. If travel cost is only a fraction of some larger cost package, then this term is expected to have a small exponent in the mathematical formulation. Therefore the proposal to alter consumption for each mode by an elasticity applied to total trip cost using only representative values of time for each mode may be an acceptable approximation.

There is an interesting consequence to altering consumption on a mode by mode basis by using the total trip cost raised to some power. Since each



mode's trip cost is calculated from some representative value of time, consumption of slow modes is insensitive to time and consumption of fast modes insensitive to cost, even if the same elasticity is used for all modes. Thus the intuitively correct behavior occurs even without adjusting the fundamental elasticity measure for different classes of consumers.

## SECTION 4

### MATHEMATICAL FORMULATION

#### 4.1 Travel Generation

Total trip generation from city  $i$  is proportional to the population  $P_i$  and to the per capita income  $I_i$ :

$$D_i \sim P_i \cdot I_i^\beta \quad (6)$$

#### 4.2 Alternative Destinations

The probability of a trip purpose not being satisfied by someone closer than city  $j$  to city  $i$  is

$$(1 - \Delta)^{L_{ij}} \quad (7)$$

where  $\Delta$  is the probability of any one person satisfying the trip purpose and  $L_{ij}$  is the population within a circle with radius  $d_{ij}$  of city  $i$ , including the population of city  $i$  itself.

The probability of the trip purpose being satisfied in city  $j$  is approximately

$$\Delta \times P_j \quad (8)$$

Combining these two probabilities gives the probability of someone in city  $j$  being the closest suitable destination:

$$D_{i \rightarrow j} \sim (1 - \Delta)^{x_{ij}} \times \Delta \times P_j \quad (9)$$

#### 4.3 Modal Split

Modal split is calculated in two steps. First the value of time at which mode  $n$  is equally costly as mode  $m$  is calculated:

$$V_{nm} = \frac{C_n - C_m}{T_m - T_n} \quad (10)$$

This expression is for  $n$ , the high speed/high cost mode. In the case of three or more modes, there is a hierarchy of these watershed values of time numbers, each one separating modes of adjacent performance.

Given this watershed value of time, the modal split for the premium mode  $n$  becomes

$$F_n = \int_{V_{nm}}^{V_{\infty}} \phi(v) dv \quad (11)$$

where  $\phi(v)$  is the probability density function for value of time  $v$ . For intermediate modes, other limits of integration are used. If  $\phi$  is assumed Gaussian in the form  $\phi(V_{avg}, V_{\sigma})$ , the integral above can be approximated by

$$\%n = \frac{1}{1 + e^{[(V_{nm} - V_{avg})1.841/V_{\sigma}]}}$$

This is merely a numerical approximation of the cumulative normal above the value  $V_{nm}$ .

#### 4.4 Price Consumption

The final part of the model states that consumption for mode  $n$  reacts to the full cost of mode  $n$ . The cost is calculated using some representative value of time  $v_n$ :

$$D_{ijn} \sim (C_n + v_n \cdot T_n)^\alpha \quad (13)$$

#### 4.5 Summary

The total demand between cities  $i$  and  $j$  on mode  $n$  is:

$$\begin{aligned} D_{ijn} = & k \cdot (C_n + v_n T_n)^\alpha \cdot P_i \cdot P_j \cdot \Delta \\ & \cdot [1 + \exp((V_{nm} - V_{avg}) \cdot 1.841/V_\sigma)]^{-1} \\ & \cdot [I_i^\beta (1 - \Delta)^{L_{ij}} + I_j^\beta (1 - \Delta)^{L_{ji}}] \end{aligned} \quad (14)$$

The following relationships appear in this model:

(1) The logit approximation of the cumulative normal, with an indication of the parameters which belong in the exponential.\*

(2) The gravity model ( $P_i \times P_j$ ) without its distance term. The

---

\* Rarely do traditional models employ a variable such as  $V_{nm}$ , but some do, perhaps without knowing exactly why. See especially Lave,(3), for an excellent intraurban example of this.

adjacent multiplications of city populations comes from two different effects -- travel generation and trip attraction. Other population information is in the terms  $\lambda_{ij}$  and  $\lambda_{ji}$ .

(3) A price-consumption relationship with a constant elasticity ( $\alpha$ ); unusual as it relates to a cost involving both money and time.

(4) Income elasticities of  $\beta$ .

This model is highly non-linear. "Choosing equations which are linear functions of the parameters (would) contribute to making the computation of the parameters a mathematically easy job. On the other hand, due to their arbitrary nature, the equations that we (would) get are useful only for summarizing the data and for interpolating between tabulated values."\*

Although calibration of non-linear equations is more difficult, their form may result in more appealing and rigorous models.

---

\*

Bard, Y., (4), p.4

## Notes for Table 5.1

CITY PAIR NAMES = from airport code of major airport in region

$P_i$  = the 1973 Bureau of Economic Analysis (BEA) region population (in millions) for the first city names in this city pair.

$P_j$  = for the second city named in the city pair

$L_{ij}$  = the 1973 populaiton (in millions) closer to city i than is city j.

This figure was derived form drawing circles on a map of BEA regions and populations and counting. The populaiton of city a is always included.

$L_{ji}$  = vice versa  $L_{ij}$ .

$I_i$  = the 1973 BEA average income for city region i.

$V_{nm}$  = the dollars per hour at which auto and air modes are equally expensive. Auto and air times\* from Part 1. Air time has 1 hour added for access and egress. Auto and air costs from Part 1 also, converted to 1973 dollars. \$6.36 added to air costs for access and egress.

$C_n$  = the 1973 air fare. Data from Part 1.

$T_n$  = the 1973 air travel time (including access and displacement times).

PAX = the 1973 average daily one-way origin-destination passengers from Part 1.

---

\* Air time includes displacement or wait time.

regression package (reference 9) was employed for the final analysis.

The calibration was largely successful. An  $R^2$  of just over 90% was achieved. However, considerable variation in most of the parameters was possible without much change in the summed squared error. The paragraphs below are a parameter-by-parameter discussion of the results.

(i)  $\beta$  -- the elasticity of consumption with average city income. The best  $\beta$  was 0.91. This compares with income\* elasticities of 1.7, 1.3 and 1.0 for long, medium and short haul air travel (Eriksen, reference 7). However, our  $\beta$  refers to short-haul travel by all modes, so a lower value is quite reasonable.  $\beta$  was largely independent of the other parameters. A broad range of values was possible, with a 1% change in standard error for  $\beta$ 's from .75 to 1.02.

(ii)  $\Delta$  -- the probability of one person being able to satisfy the trip purpose. The best  $\Delta$  was  $0.018/10^6$ . There was little guidance for a realistic value of  $\Delta$ , but the number was not unreasonable. A modest variation produced no more than a 1% change in summed squared error. The range was .0172 to .0207.  $\Delta$  was largely independent of the other parameters.

(iii)  $\alpha$  -- elasticity of air travel consumption with cost. Cost in this case was the sum of ticket price (1973 dollars), \$6.36 in access cost, and \$7 times travel time, including average displacement times,\*\* access, and

---

\* Income was defined by a more complicated measure by Eriksen; however, the results should be comparable.

\*\* Average displacement was developed by Eriksen (reference 7) and Part 1 for this data. This is the average difference between desired departure time and the closest departure available.

aircraft time,  $\alpha$  was  $-.69$ . (For a 1% change in error, a range from  $-.5$  to  $-.8$  was possible.) This translates approximately to a level of service exponent of  $-0.27$ . Eriksen found level of service exponents for long and medium haul travel of  $-.49$  and  $-.70$ , respectively. The reduced sensitivity to time violates the progression observed by Eriksen. Coupled with the broad range of reasonably satisfactory values and the doubt about the value of time, this throws the observed  $\alpha$  into some doubt.

(iv) value of time -- \$7 (1973 dollars). While value of time was not a regression parameter, experimental variations in this value were made. There was a trend to high values of time, eliminating the effect of fare from the model as observed in Part 1. However, the effect of reducing error by increasing the value of time was mild, so the use of the arbitrary value was continued. \$7/hour was the mean household income in the period studied, assuming a 2000 hour year.

(v)  $V_{avg}$  -- the average value of time among the traveling public assuming a Gaussian distribution of such values.  $V_{avg}$  was forbidden to go below \$0. The "best" value was \$.04/hour, but \$0.0 values were possible with almost no change in assumed squared error. As discussed earlier, the Gaussian distribution was used as a preliminary assumption and is undoubtedly a poor one. The statistical preference for low  $V_{avg}$  values was an attempt to employ only the right half of the Gaussian to model the income distribution.

(vi)  $V_{\sigma}$  -- the standard deviation of the Gaussian distribution of values of time. The best value was \$6.26. This was a fairly stable value and may provide some guidance in establishing a more reasonable curve shape. The value itself is not at all unreasonable.



(vii) k -- scale factor for all demands

Demand was measured in one way origin-destination air passengers per day. With time in hours, populations in millions, and money in 1974 dollars, k was 16.3.

Discussion of Statistical Accuracy

The model employed a non-linear formulation and thus the usual statistical tests were not possible. Work in Part 1 using the same data base revealed no serious statistical problems. The minimum summed squared error was not a very distinct minimum, so the results should be taken as establishing reasonable ranges for the parameter values.

The model was subjected to one statistical test. A test for the stability of the parameters was made by removing first the 5 largest markets and then the 5 smallest markets from the data set and running new calibrations. The results in Table 5.2 suggests the estimates are not unduly biased by large or small observations. The only parameter of doubt is  $\Delta$ , which is much larger without the dense markets. There is apparently a tendency for travel between large cities to be not sufficiently explained by other parts of the model. However, large markets have higher levels of service, so most likely the error is in either the demand distribution at very low values of time or in  $\alpha$ , the sensitivity to levels of service. Indeed, the statistical favoring of higher values of time than the one employed suggest the same conclusion.

## SECTION 5.

## CALIBRATION

The research which led to this model formulation was concerned with predicting short haul air travel, so the model used air travel data. Air passenger travel in 33 city-pair markets during 1973 was used. The data were true origin-destination data reported by the airlines\* and included some passengers connecting to international flights. Travel time and costs were obtained from the Official Airline Guide(6). Travel time included a displacement time calculated from the observed schedule and an assumed distribution of ideal departure times(7). The list of city-pairs did not include resort towns or any low density service air markets.\*\* Distances were between 100 and 400 miles. Air access times and costs as well as auto, bus and rail times and costs were calculated. The mode adjacent to air was always auto.

Using the 1973 BEA incomes and populations from the Commerce Department, the predicted and actual city-pairs demands were compared (data in Table 5.1). Six parameters ( $\beta$ ,  $\Delta$ ,  $\alpha$ ,  $V_{avg}$ ,  $V_{mode}$ , and a scale constant) were adjusted to minimize the summed squared error.

---

\*Source: CAB O-D Table 10 (5). Passengers connecting to international flights are included in the data, but no on-line or off-line connections to domestic markets are included.

\*\*Where frequency is low, average air travel times fall below car times and distributions about the average become important.

Table 5.1 Data

city pair	$P_i$	$P_j$	$\lambda_{ij}$	$\lambda_{ji}$	$I_i$	$I_j$	$V_{nm}$	$C_n$	$T_n$	PAX
MKC OMS	2.3	0.8	3.0	3.0	5072.	5111.	7.8	25.53	2.35	87.3
DFW OKC	2.8	1.2	6.0	3.7	5064.	4385.	6.8	26.36	2.21	145.4
DAY DTT	1.2	5.2	18.0	15.0	5121.	6089.	9.2	26.36	2.97	52.4
PIT WAS	3.7	3.2	18.0	23.0	4700.	6358.	4.4	27.19	2.40	196.3
DTT PIT	5.2	3.7	20.0	18.0	6089.	4700.	3.4	27.19	2.60	144.0
CLF DTT	4.2	5.2	6.1	6.3	5304.	6089.	7.6	22.20	1.94	221.3
FAR MSP	0.3	3.0	1.0	6.0	5609.	5216.	6.9	28.04	2.86	56.8
DAY PIT	1.2	3.7	18.0	19.0	5121.	4700.	3.8	28.04	2.97	47.0
ORE PHL	1.3	7.4	13.0	32.0	4475.	5310.	1.7	28.04	2.60	99.4
CVG RNA	1.9	1.5	19.0	19.0	4858.	4177.	6.1	28.87	3.23	35.9
CVG PIT	1.9	3.7	32.0	29.0	4858.	4700.	3.5	29.70	2.80	58.0
MFM STL	1.8	3.2	10.0	35.0	4838.	4907.	1.2	29.70	2.44	96.5
MKF MSP	2.1	3.0	28.0	9.5	5414.	5216.	2.3	33.87	3.06	139.7
DFW JAN	2.8	1.1	14.0	21.0	5064.	4367.	3.0	38.87	3.43	48.4
DEN SLC	1.7	1.2	3.5	2.0	5420.	4227.	0.4	38.04	2.60	200.3
DAY STL	1.2	3.2	51.0	45.0	5121.	4907.	4.3	35.53	2.91	45.1
BOS ROC	6.5	1.0	41.0	51.0	5227.	5534.	3.5	38.87	2.87	126.4
MEM MSY	1.8	2.2	27.0	13.0	4838.	4062.	4.2	36.37	2.63	96.5
HOM MSY	2.5	2.2	13.0	9.9	5099.	4062.	2.1	33.87	2.30	326.9
ALB PIT	1.4	3.7	53.0	72.0	4833.	4700.	2.0	41.37	3.91	35.9
ATL MEM	2.5	1.8	20.0	18.0	4838.	4838.	1.2	35.53	2.62	179.4
ATL CVG	2.5	1.9	53.0	26.0	4838.	4858.	1.3	37.21	2.78	82.2
SAN TUS	1.5	0.5	21.0	3.4	5204.	4550.	3.9	37.21	5.95	40.8
NYC WAS	18.2	3.2	43.0	26.0	6115.	6368.	3.6	28.04	2.05	1012.0
CHI STL	8.2	3.2	30.0	29.0	5983.	4907.	3.4	29.70	2.18	620.0
CHI DIT	8.2	5.2	23.0	29.0	5983.	6089.	3.8	28.87	2.21	828.5
BOS PHI	6.5	7.4	34.0	49.0	5227.	5310.	3.3	33.3	2.48	528.3
NYC PIT	18.2	3.7	54.0	56.0	6115.	4700.	2.1	35.53	2.36	836.9

Table 5.2 Variation in Parameters

Parameter	Value	1% Range
$\beta$	.91	0.71 to 1.02
$\Delta$	.018	.0172 to .0207
$\alpha$	-.69	-.50 to -.80
$V_{avg}$	\$0.04	-3.0 to 1.3
$V_{\sigma}$	\$6.26	5.8 to 7.1

Table 5.3 Tests for Robust Estimates

	Original	Less Top	Less Bottom
	<u>Data</u>	<u>5 Markets</u>	<u>5 Markets</u>
Points	33	28	28
$\beta$ (income)	.91	.88	.90
$\alpha$ (price)	-.69	-.61	-.69
$\Delta$ (people)	.018	.050	.018
$V_{avg}$	.04*	.0*	.0*
$V_{\sigma}$	\$6.26	\$5.11	\$6.19

\* constrained to be  $\geq 0$

## SECTION 6.

## Conclusions

The model developed for daily, one way origin-destination air travel in a city pair is given by:

$$D_{ijn} = k \cdot \Delta \cdot P_i \cdot P_j \cdot (C_n + v_n \times T_n)^\alpha / [1 + \exp((V_{nm} - V_{avg}) \cdot 1.841/V_\sigma)] \\ \times \{(1 - \Delta)^{\lambda_{ij}} \cdot I_i^\beta + (1 - \Delta)^{\lambda_{ji}} \cdot I_j^\beta\}$$

where

$D_{ijn}$  = the daily one-way origin-destination travel in the city-pair  $ij$  on air mode  $n$

$k$  = 16.33 one-way passengers per day

$\Delta$  = 0.018 trip destinations per million of population

$P_i, P_j$  = the population (in millions) of cities  $i$  and  $j$

$C_n$  = the cost of the air trip, in 1974 dollars, including access

$T_n$  = the time of the air trip, including access and displacement time, in \$/hour

$v_n$  = the value of time = \$7/hr in 1974 dollars

$\exp$  = the exponential function

$V_{nm}$  = the value of time at which car travel has the same cost in time and money as air travel

$V_{avg}$  = \$0.04/hour the center of a normal distribution of value of time, only the right hand side of which is valid.

$V_\sigma$  = \$6.26/Hour the standard deviation of the distribution of value

$\lambda_{ij}$  = the population(in millions) living within a circle centered at

city  $i$  which just misses city  $j$ . This value includes the population of city  $i$

$\ell_{ji}$  = vice versa

$I_i, I_j$  = the average annual per capita incomes (1973 dollars) for cities  $i$  and  $j$

$\alpha$  =  $-.69$ , trip consumption elasticity with respect to total of time and money cost

$\beta$  =  $.91$ , elasticity with respect to income

This formula is a rearrangement of four multiplicative effects:

(i)  $K \cdot P_i I_i^\beta$  -- the trip generation from city  $i$  to all destinations.

(ii)  $(1 - \Delta)^{\ell_{ij}} P_j \cdot \Delta$  ---- the probability of no person closer than city  $j$  satisfying the trip purpose times the probability of someone at  $j$  satisfying the trip purpose. This term takes the place of previous distance terms and measures the effect of intervening opportunities on demand.

(iii)  $1/[1 + \exp\{(V_{nm} - V_{avg}) \cdot 1.841/V_\sigma\}]$  -- the modal split term.

This is a numerical approximation of the cumulative population with a value of time above that necessary to make the air mode preferable to the next slower and cheaper mode.

(iv)  $(C_n + V_n \times T_n)^\alpha$  -- the elasticity with respect to total trip cost including the cost of the time necessary to make the trip.

This formula is logically derived from behavioral assumptions and calibrated from a small sample of 33 city pairs with reasonable competition between air and auto modes ( $\$0 < V_{nm} < \$10$ ) and no great tourist attractions. It includes the effects of intervening opportunities on diluting demand in a city pair, the effects of modal split, the travel

generating effects of income and population, and the travel discouraging effects of air travel price and time.

Alternative modes are captured in the definition of  $V_{nm}$ , the value of time above which the air journey is cheaper in total cost. When a mode such as high speed ground or unconventional air is introduced, the competing mode is the next slowest and cheapest below air (be it car or rail). If a faster, more expensive mode is added to the market, the formula predicts the modal traffic for the combined air and super-air modes. A second modal split can be made without further calibration according to the logic of section 3.2.

The particular calibration of this model cannot be said to be "good". Considerable ranges in the parameters  $\gamma$ ,  $\beta$ ,  $\Delta$ ,  $V_n$ , and  $V_{avg}$  are all reasonable.

This model blends gracefully into long haul air travel models as the modal split term gradually allocates all travel to the air mode. As distances increase, the intervening opportunities terms  $(1 - \Delta)^{2_{ij}}$  become more consistent among cities and change less with distance.\*

---

\* For medium and long haul travel the assumption that level of service is relatively independent of O-D demand fails and a two equation model should be employed, as in reference 1.



## SECTION 7

## FUTURE WORK

The early results from the behavioral model suggest that it is worthwhile to continue this investigation along the same lines. Two of the advantages of a behavioral formulation are that components of the model can be calibrated from very modest survey data and any improvement of any component allows more meaningful calibration of all the other components, even using the old data. At this point, improvements in mathematical form and accuracy are both relatively easy to map out. A few specific improvements can be suggested:

- (i) The logit form for modal split does not appear justified.

Figure 7.1b shows the modal split vs. distance relationship for 17 Northeast Corridor markets. This was the formulation suggested in Part 1. Figure 7.1a shows the critical value of time vs modal split for the same city-pairs.

The behavioral model appears as accurate as the distance model and much more logical. However, the classic bell curve does not appear a correct description of the distribution of value of time. In Part 2, parameters were adjusted to produce the curve in Figure 7.2. For comparison, 1973 income tax returns produced the hourly income distribution in Figure 7.3a. If half the trips are business trips (at 2x hourly wage) and half pleasure trips (at 1/2 hourly wage) the value of time distribution follows Figure 7.3b.

One problem is that data on air trips vs income show the tails of the income distribution absorb a good part of the travel at least in the long haul. This makes such constructive efforts as above difficult.

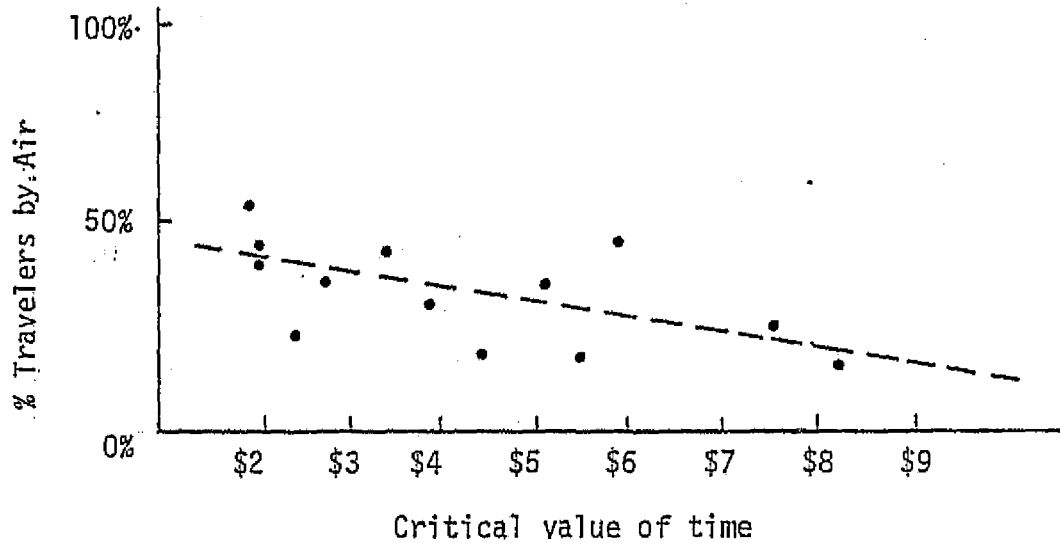
However, work should proceed along three lines:

- (1) Improved income and propensity to travel distributions. Data from on-board surveys by airlines, New York Port Authority, and the Northeast

Modal split from 1968 Northeast Corridor Data

Figure

7.1a Fraction of travelers by air vs value of time for equal cost air and auto journeys.



Figure

7.1b Fraction of travelers by air vs intercity distance, same city pairs

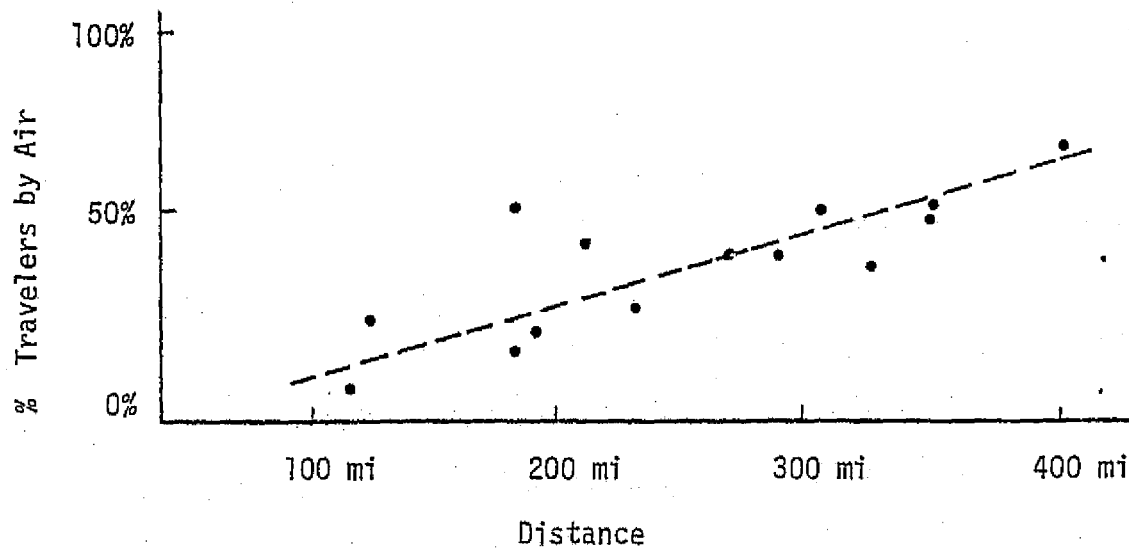


Figure 7.2: Value of time distribution as indicated by data

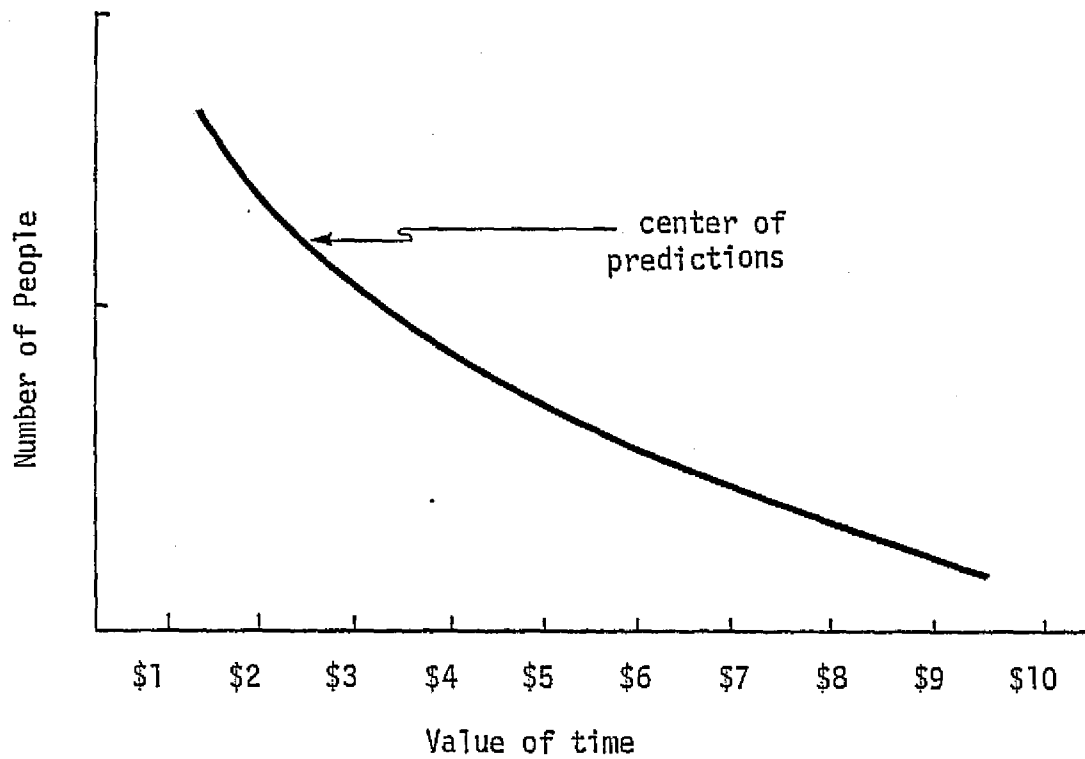
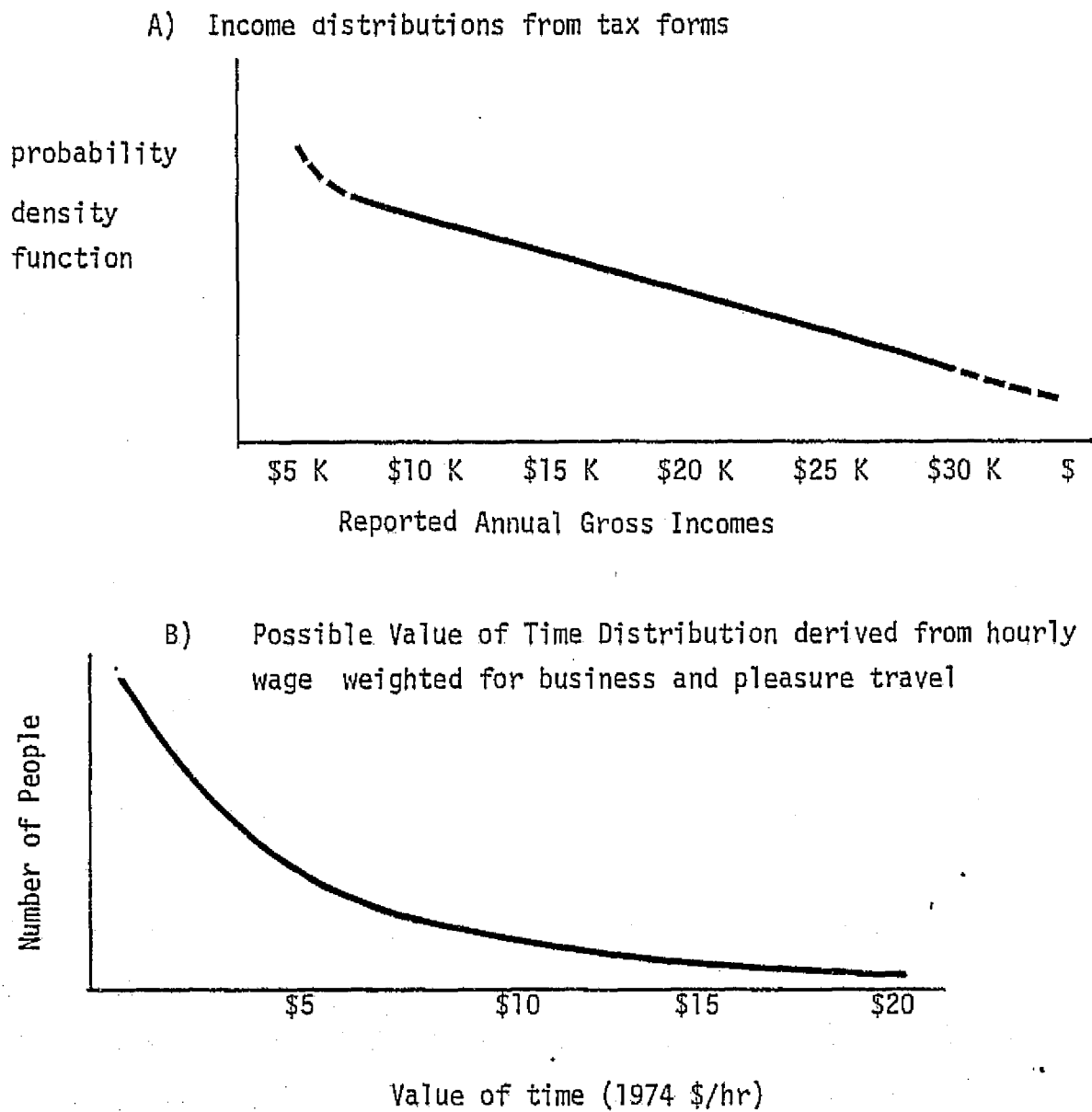


Figure 7.3: Possible Value of Time Distributions

Corridor study may lead to deriving suitable mathematical forms.

(2) Statistical and analytical studies of special markets where the air to car mode performances are unusual.

(3) Use of local income figures to determine means and variances of value of time distribution.

(ii) Consumption elasticity with respect to trip cost ( $\beta$ ) may well not be a constant. In fact  $\beta$  has a distribution among consumers which correlates with their value of time. Behavioral modeling of demands using Monte Carlo simulations may allow the development of a mathematical form simple enough for use in such models and more suited to reasonable behavior patterns. With knowledge of the distribution of incomes (about the average urban income) and knowledge of the propensity to travel, both the value of time constant in the air travel cost term and the demand dependence on cost should be better predicted.

(iii) The income elasticity ( $\alpha$ ) formulation can be studied and reformulated in the same way as cost elasticity,  $\beta$ , above.

(iv) City populations may be an adequate measure of the propensity to travel, but they are an inadequate measure of the probability of trip purposes being satisfied in a city. Cities like Las Vegas or Miami attract two to three times the travel their populations indicate in our model, while industrial cities attract no more than half. It is quite possible to study a table of city-pair demands and establish a population adjustment factor for each city in the table. (This is possible because the data points go up as  $n^2$  and the unknowns as  $n$ , where  $n$  is the number of cities involved.) With this accomplished, it may be possible to

predict this population adjustment factor from data on retail, manufacturing, or service industry activities for the city. While this is a major effort, it is the only hopeful approach to demand prediction without place-specific adjustment factors. The problem is that pleasure travel is to a great extent place-specific and even business demand responds to historical locations of trade which are also place-specific. A good deal of a city's specific attractiveness factor may be a coincidence rather than logically explained.

(v) There is undoubtedly a small change in  $\Delta$ , the probability of a person being a useful trip destination, with total U.S. population.  $\Delta$  should have declined slightly through the years. Work on historical trends in this dimension should be considered.

(vi) Access and egress times should be modeled in detail and Monte Carlo studies made of the effects of distribution of access and wait time. Particularly in the 100-200 mile range such considerations appear to dominate. The current use of approximate averages should be continued only after it is verified by more detailed studies.

(vii) The definition of intervening population employed in this study was all people closer in miles than the destination city. In practice, there should be two adjustments to this. First, people should be closer in travel time and cost, not miles. This puts a well-served city 400 miles away closer than an ill-served 300-mile neighbor. Second, the presence of large populations at the same distance as the destination city should exert a diluting influence on demand.

The best way to examine these problems at the moment is to do a simulation of travel for the sample city-pairs. The purpose of the

simulation would be to see if the problems are important or whether they can be expected either to average out or to be expressed by a simple functional curve shape. The simulation would involve generating a person, assigning values of time, locations, destinations, and departure times to him according to representative probability distributions, and choosing his mode of travel. Repetition of this process until a stable cumulative pattern is observed should be very useful.

This procedure is also one of the few ways of drawing up statistical tests for non-linear regression models.

## REFERENCES - Part 1

1. Alcala, Roger E., An Economic Study of the Demand for Air Passenger Service over the North Atlantic, Ph.D. Thesis, Princeton University, University Microfilms, Inc., Ann Arbor, Michigan, 1969.
2. Ibid., pp. 191-208
3. Ibid., p. 140
4. All American Train Fares ----- Commerce Passenger Fares, AMTRAK (National Railroad Passenger Corporation), Washington, D.C., Edition 2, August 16, 1972; Edition 3, July, 1973; Edition 4, February 1, 1974.
5. Blumer, Terry P., "A Short Haul Passenger Demand Model for Air Transportation," M.S. Thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology (September 1976).
6. Bureau of the Census, National Travel Survey, 1972 Census of Transportation, U.S. Department of Commerce, Washington, D.C. 1973.
7. Bureau of the Census, National Travel Survey Public Use Tape User's Manual, 1972 Census of Transportation, U.S. Department of Commerce, Washington, D.C., September 1973.
8. Bureau of the Census, Annual Survey of Manufacturers, U.S. Department of Commerce, Washington, D.C.
9. Bureau of the Census, Current Population Reports, U.S. Department of Commerce, Washington, D.C.
10. Caves, Richard E., Air Transport and Its Regulators, Harvard University Press, Cambridge, Massachusetts, 1962.
11. Charles River Associates, A Disaggregated Behavior Model of Urban Travel Demand, U.S. DOT Contract No. FH-11-7566, Final Report, March 1972.
12. Chow, G.C., "Tests of Equality Between Sets of Coefficients in Two Linear Regressions", Econometrica, Volume 28, July 1960.
13. Origin-Destination Survey, Table 8, Civil Aeronautics Board, Washington, D.C.
14. Clarke, G.T., and Schkade, L.L., Statistical Methods for Business Decisions, South-Western Publishing Company, Cincinnati, Ohio, 1969.



15. Crow, Robert T., Young, K., and Cooley, T., "Alternative Demand Functions for 'Abstract Mode' Transportation Models", Transportation Research, Volume 7, pp. 335-354.
16. Eriksen, S.E., and Taneja, N.K., "Unified Performance Measures of Quality of Airline Service Between Two Regions Encompassing Several Difficult to Quantify Scheduling Variables", M.I.T. Department of Aeronautics and Astronautics. Presented November 17, 1975 at the Joint National Meeting of the Operations Research Society of America and the Institute of Management Sciences, Las Vegas, Nevada.
17. Eriksen, S., Scalea, J.C., and Taneja, N.K., Foundations of a Methodolgy for Determining the Relationship Between Supply of and Demand for Air Transportation Services, M.I.T. Department of Aeronautics and Astronautics, Contract No. NAS2-8157 for Ames Research Center, NASA.
18. Ibid., p. 38.
19. Ibid., p. 36.
20. Fan, S.L., Horonjeff, R., Kanafani, A., and Mogharabi, A., Forecasting the Demand Potential for STOL Air Transportation, prepared for Ames Research Center, National Aeronautics and Space Administration, February 1973.
21. Hollander, Stanley C., "Demand for Intercity Passenger Travel in the Washington-Boston Corridor", Systems Analysis & Research Corporation paper in Passenger Transportation, MSU Business Studies, Michigan State University, East Lansing, Michigan, 1968.
22. Massachusetts Institute of Technology, The Impact of Changing Technology on the Demand for Air Transportation, Semiannual Status Report, M.I.T. Flight Transportation Laboratory, May 3, 1976.
23. Neter, J., and Wasserman, W., Applied Linear Statistical Models, Richard D. Irwin, Inc., Homewood, Illinois, 1974.
24. Official Airline Guide, North American Edition, Reuben H. Donnelley Corporation, Oak Brook, Illinois.
25. The Official Guide of the Railways, National Railway Publication Company, New York, New York, Volume 1, Number 2, September 1974.
26. The Official Guide of the Railways, National Railway Publication Company, Volume 1, Number 2, September 1974.
27. Road Atlas, Rand McNally, Chicago, Illinois, 50th Edition, 1974.
28. Sloss, James and Kneafsey, James T., "The Demand for Intercity Rail Passenger Service", Transportation Journal, forthcoming.

29. Stuart, M.M., The Use of Gravity Model in Interpreting Air Passenger Traffic Between Hawaii and the Coterminous United States, University Microfilms, Ann Arbor, Michigan, 1968.
30. Survey of Buying Power, Sales Management, New York, New York.
31. Ibid., p. C-3.
32. U.S. Department of Commerce, Business Conditions Digest, No.'s 140 and 145, Table 1: "Summary of Recent Data and Current Changes for Principal Indicators".
33. United Technologies Research Center, "Transportation Demand Forecasting," Report R76-111240-1, for NASA, January 1976.
34. Ibid., pp. 38-39.
35. Ibid., p. 4
36. Ibid., pp. 17-18.
37. Ibid., p. 10.
38. Verleger, Philip K., A Point-to-Point Model of Demand for Air Transportation, Ph.D. Thesis, M.I.T. Department of Economics, June 1971.
39. Ibid., pp. 80-81.
40. Ibid., pp. 77-83.
41. Ibid., pp. 22-24 and 93-94.

## REFERENCES-PART II

1. Larson, S. and Odoni, A., Urban Operations Research (in preparation).
2. Morton, Alexander L., Freight Demand. Ph.D. Thesis, Harvard, 1973.
3. Lave, Charles A., "A Behavioral Approach to Modal Split Forecasting", Transportation Research, Vol. 3, pp. 483-486, 1969.
4. Bard, Y., Nonlinear Parameter Estimation, Academic Press, 1974.
5. Civil Aeronautics Board, Origin and Destination Survey 1973, Table 8.
6. The Official Airline Guide, published monthly.
7. Eriksen, Steven , Policy Oriented Multi-Equation Models of U.S. Domestic Air Passenger Markets. M.I.T. Ph.D. Thesis, 1977.
8. Civil Aeronautics Board, Domestic Passenger-Fare Investigation, US GPO 003-006-00071-9, 1976.
9. Biomedical Computer Programs, UCLA Press, 1975.

APPENDIX A  
SOURCES OF DATA

The specific sources of empirical information for the short haul demand model are displayed in Figure A.1. A set of 60 city-pairs were initially selected for the sample. This was later reduced to 49 city-pairs to avoid data collection difficulties. Annual interstate data were originally collected for a three year period (1972-1974) and pooled into a single data base.\* The sampling criteria are described as follows: The individual cities were selected to provide adequate geographical coverage of the continental U.S. and represent cities both large and small. The geographic coverage avoids biases in the regression coefficients one would expect in a regional sample. It is common, particularly in short haul models, to see a sample focused on California, or the Northeast Corridor, or New York. None of these can legitimately represent the whole U.S. In addition, the sample density by region varies roughly with the population density. So more sample cities appear in the East than in the South and West. Finally, the cities were ranked by market size (total disposable personal income) using five levels, where 1 represents a small city such as Fargo, North Dakota, and 5 represents a large city such as New York (see Figure A.2.

Once a sample is selected, the city boundaries must be defined. Here

---

\*The actual city pair selection was designed by Steven Eriksen in a short haul sample as part of a much larger data base assembled for related work being performed at M.I.T. (17). Eventually, the sample data were reduced to cover the years 1972 and 1973 only, as explained below in Section 2.4.4

Figure A.1: Sources of Data for Short Haul Demand Model

Data	Source	By
Air Trips	Origin-Destination	Civil Aeronautics Board
Effective Buying	Survey of Buying	Sales and Marketing Management
Distance : Air	Official Airline Guide	Reuben H. Donnelley Corp.
: Auto	Road Atlas Guide	Rand McNally
Air Schedules	Official Airline Guide	Reuben H. Donnelley Corp.
Rail Schedules	Official Guide of the Railroads	National Railway Publication Co.
Air Fare	Official Airline Guide	Reuben H. Donnelley
Rail Fare	All American Train Fares	AMTRAK
Auto Travel Time	"Road Atlas"	Rand McNally

Figure A.2: Sample cities organized by size where 1 is small and  
5 is large.

1

PWM (Portland, ME)  
ERI (Erie, Penn)  
TYS (Knoxville, TN)  
LEX (Lexington, VA)  
FAR (Fargo, ND)  
LBB (Lubbock, TX)  
JAN (Jackson, MS)  
RNO (Reno, NV)  
LAS (Las Vegas, NV)  
TUS (Tuscon, AR)

2

ROC (Rochester, NY)  
RIC (Richmond, VA)  
ORF (Norfolk, VA)  
DAY (Dayton, OH)  
OMA (Omaha, NB)  
ICT (Wichita, KS)  
OKC (Oklahoma City, OK)  
SLC (Salt Lake City, UT)  
SAC (Sacramento, CA)

3

ALB (Albany, NY)  
RD (Raleigh, NC)  
MEM (Memphis, TN)  
BNA (Nashville, TN)  
CVG (Cincinnati, OH)  
MKE (Milwaukee, WI)  
MSY (New Orleans, LA)  
DEN (Denver, CO)  
SAN (San Diego, CA)

4

ATL (Atlanta, GA)  
PIT (Pittsburgh, PA)  
MSP (Minneapolis-St. Paul, MN)  
MKC (Kansas City, MO)  
STL (Saint Louis, MO)  
DFW (Dallas-Fort Worth, TX)  
HOU (Houston, TX)

5

BOS (Boston, MA)  
NYC (New York, NY)  
PHL (Philadelphia, PA)  
WAS (Washington, D.C.)  
CLE (Cleveland, IL)  
DTT (Detroit, MI)  
CHI (Chicago, IL)  
LOS (Los Angeles, CA)

the Metropolitan County Areas (MCA's) are used, as defined in the 1969 issue of the Survey of Buying Power (30). These areas are equivalent to the Census Bureau's Standard Metropolitan Statistical Areas everywhere except in New England. These areas appear to give proper coverage for a short haul air market because they are small enough to imply short access times, an important point in a short haul market. The BEA regions, for example, would be much too large. Few persons would drive for 100 miles to fly another 100 miles. On the other hand, the MCA's are large enough to capture most travelers, because they include the surrounding suburbs as well as the central city.

The city-pairs, like the individual cities, also should reflect good geographical coverage as well as sufficient market size coverage. The market size rankings produced fifteen combinations when the cities were paired (1-1, 1-2, ....2-3....5-5). Eriksen chose to draw four samples from each of the 15 combinations, yielding sixty city pairs that represented sufficient geographic coverage of the U.S. For this study, however, the sample was reduced to 49 because of data collection difficulties.

## APPENDIX B

Empirical Support for the Short Haul Demand ModelB.1 Mass

The gravity term in the short haul demand model requires that each city be represented by its mass, where the mass of a city captures its ability to generate and attract travelers. The income is used for mass in this study, but by no means is the only way to capture mass. Several methods have been tried and a couple of interesting mass variables will be discussed here.

B.1.1 Income Versus Population

Population and income are the two most common mass variables. Population received much attention in early demand modeling efforts, but income has replaced it as the most frequently used mass variable. Income is thought to be a better measure of the ability of persons to travel. Alcaly explains why:

The aggregate purchasing power of the consumers residing at a given node is clearly superior to the node's population as a measure of its traffic generation potential. For, if the total income at node  $i$  were very low (near the subsistence level), no travel could result regardless of the number of people living there. Similarly, a node's attractiveness as a destination is better represented by its income than its population. The availability of "amenities" such as comfortable accommodations, good internal transportation, etc., is undoubtedly more closely related to an area's income than its population....(3)

Income not only works moderately well but is easier to collect and use. A number of sources have income related data and the only modification necessary is an adjustment for inflation if time-series data are used. However, income is only a very simple variable and cannot possibly account for all of the



attractions occurring between cities. More complex mass variables are needed; variables that incorporate more of the attraction phenomena occurring between cities.

### B.1.2 Income Distribution

It might be said that neither population nor income truly measures the ability of a city to generate travelers. Population is poor because a large city would still not produce many travelers if all residents were poor. Income also has the same propensity to generate a trip -- a man who is twice as rich will take twice as many flights. Surveys taken indicate that flight generation is not linear with income, but increases at an increasing rate (for the income levels sampled). Some surveys contained in Verlager's dissertation support this hypothesis. He suggest that the number of trips per household increases in a 1, 1 1/2, 4 1/2, 16 progression for income levels \$0-4999, \$5000-9999, \$10,000-14,999,  $\geq$  \$15,000, respectively--a decidedly non-linear progression. Verlager developed a variable equal to the weighted sum of households by income level, weighted by propensity to travel.

$$M_i = \sum_{r=1}^R N_i(r) H_i(r) \quad (B.1)$$

$M_i$  = Mass of city i

$N_i(r)$  = number of households in city i in income level r

$H_i(r)$  = Average number of air trips taken by a household from city i  
in income level r

This variable should indicate the true propensity of a city to generate personal and pleasure type traffic. It cannot claim to measure business

traffic , though close correlation undoubtedly exists.

The above specification builds a mass variable for a single equation, gravity model, but what about a modal split model? A mass variable capable of capturing travel by different modes and by income level requires extension of the model proposed by Verleger. A separate mass variable can be specified for each mode using a specification similar to equation (B.1). One suggestion is to simply sum over each mode. Several sources contain trip data by income level including the National Travel Survey, and the New York Port Authority (see Tables B.1 and B.2). An example is shown below.

$$M_{im} = \sum_{r=1}^R N_{i(r)} H_{i(m)} \quad (B.2)$$

$M_{im}$  = Mass of city i given a mode m

R = Number of predefined income levels

$N_{i(r)}$  = Number of households from city i in income level r

$H_{i(r)m}$  = Average number of trips per household from city i in income level r using mode m

$$M_i = \sum_{m=1}^K M_{im} = \sum_{m=1}^K \sum_{r=1}^R N_{i(r)} H_{i(r)m} \quad (B.3)$$

K = Number of modes

### B.1.3 Generation-Attraction Model

None of the aforementioned mass variables attempt to describe the attraction process in detail. The following discussion attempts to lay a foundation from which a comprehensive mass variable can be built.

Consider two cities, city i and city j, and the traffic originating in city i and destined for city j. A force exists between the two cities

\* Table B.1

## Relationship Between Air Trips and Family Income, 1964 and 1967

<u>Income Range</u> <u>(000 omitted)</u>	<u>Percent of</u> <u>Respondents</u>	<u>Percent of Households</u> <u>in Income Group</u>	<u>= Air Trips Per</u> <u>Household</u>
---	---	--	--

## 1963 Port Authority of New York Survey

0 - 3 (dollars)	2	25.7	.08
3 - 5	4	17.5	.23
5 - 6	4	10.3	.39
5 - 7	4	9.3	.43
7 - 10	11	19.8	.55
10 - 15	25	12.6	1.98
15 - 20	16	----	----
20 - 25	9	3.8	6.58
25 - above	25	.8	31.10

## 1963 U.S. Census of Transportation

0 - 1	1	7.2	.10
1 - 2	2	10.0	.20
2 - 3	1	8.5	.10
3 - 4	2	8.7	.20
4 - 5	5	8.8	.60
5 - 6	5	10.3	.50
6 - 7	7	9.3	.70
7 - 10	12	17.8	.70
10 - 15	30	12.6	2.40
Above 15	24	4.7	5.10

## 1967 Port Authority of New York Survey

0 - 5	7	33.7	.28
5 - 7.5	9	19.8	.45
7.5 - 10	9	17.0	.53
10 - 14	21	19.1	1.11
15 - 20	17	----	----
20 - 25	11	8.1	4.59
25 - 50	17	----	----
Above	9	2.1	12.38

(Table B.1, continued)

<u>Income Range</u> <u>(000 omitted)</u>	<u>Percent of</u> <u>Respondents</u>	<u>Percent of Households</u> <u>in Income Group</u>	<u>= Air Trips Per</u> <u>Household</u>
1967 U.S. Census of Transportation			
0 - 2	1.4	13.0	.11
2 - 3	1.3	7.2	.18
3 - 4	2.3	6.9	.33
4 - 5	2.2	6.6	.33
5 - 6	4.5	7.6	.59
6 - 7.5	8.2	12.2	.65
7.5 - 10	14.6	16.2	.88
10 - 15	27.4	19.1	1.43
15 - 25	18.8	8.1	2.32
25 + above	9.4	2.1	4.48

Sources: Port Authority of New York, "New York's Domestic Air Passenger Market," April, 1963 through March, 1964, (Published May, 1965) and April, 1966 through March, 1967 (Published May, 1969).

Published by Port Authority of New York.

U.S. Bureau of the Census, Census of Transportation, 1963,

(Vol. 1), Passenger Transport Survey, p. 18, Washington, D.C.

U.S. Government Printing Office, 1965; National Travel Survey,

1967, (Vol. 1), Passenger Transport Survey, p. 23, Washington,

D.C., U.S. Government Printing Office 1969; Current Population

Reports, Series P-60, No. 66, "Income in 1968 of Families and

Persons in the U.S.A.," Government Printing Office, Washington, D.C., 1969, p. 20.

\*Verleger, (38), pages 80-81.

Table B.2: Number of Trips Per 100 Adults for Nonbusiness Purposes,  
by Income Group and Mode of Transportation, 1956

Income Class	Mode of Transportation			
	Air	Rail	Bus	Private Auto
Under \$1,000	1.7	6.0	8.5	37.8
1,000-1,999	1.4	9.3	14.9	60.7
2,000-2,999	2.4	13.6	7.9	91.3
3,000-3,999	3.0	9.7	8.1	130.8
4,000-4,999	3.4	12.0	5.8	126.4
5,000-5,999	5.8	13.8	8.9	192.6
6,000-7,499	11.6	10.4	4.8	194.3
7,500-9,999	21.2	13.8	9.8	271.1
10,000-14,999	21.2	13.9	8.1	295.7
15,000-19,999	54.7	17.2	5.6	276.5
20,000 and up	76.2	38.4	13.1	201.1

Source: CAB, Bureau of Air Operations, Bureau Counsel Exhibit No. BCR-188  
(Sup.), General Passenger Fare Investigation, Table 11, p.18

\*Reference (10)

which produces the traffic flow. This traffic is the combined result of city  $i$ 's ability to generate traffic and city  $j$ 's ability to attract traffic, i.e., the force between two cities is directly proportional to  $G_i$  and  $A_j$  where  $G_i$  equals the generating power of city  $i$  and  $A_j$  equals the attracting power of city  $j$ . Furthermore, consider the traffic to be composed of four independent groups: visiting friends and relatives (VFR), personal (PE), pleasure (PL), and business (B). Each group generates its own force and therefore has an independent attraction, i.e.

$$\begin{aligned}
 G_i * A_j &= G_{VFRi} * A_{VFRj} \\
 &+ G_{PEi} * A_{PEj} \\
 &+ G_{PLi} * A_{PLj} \\
 &+ G_{Bi} * A_{Bj}
 \end{aligned}
 \tag{B.4}$$

So the total attraction equals the sum of the attractions by travel type. This formulation assumes each trip type is independent, a reasonable assumption unless trips have mixed purposes. However, for now let us assume complete independence. The next step then is to specify each generation and attraction function by trip type.

#### B.1.3a Visiting Friends and Relatives (VFR)

Visiting friends and relatives accounts for substantial traffic. The functions for generation and attraction appear quite simple. To travel, a person requires money, so income and income distribution appear to be likely

candidates. Referring back to the discussion in Section B.1.2, income distribution (ID) is the best explanatory variable for traffic generation. Attraction, on the other hand, is quite different. A VFR traveler intends to visit people, be attracted by people, so the measure of attraction should measure people, i.e. population.

$$G_{VFRi} = ID_i = \sum_{r=1}^R N_i(r) H_i(r) \quad (B.5)$$

$$A_{VFRj} = P_j \quad (B.6)$$

### B.1.3b Personal (PE)

The mass variables now reflect a more typical construction. The ability to generate trips again corresponds to the individual's ability to travel, therefore the generating mass equals the income distribution. The attractions are not as clear; why does a personal traveler travel? It could be people, or school, or related to a unique service offered in the destination city (e.g., medical). Alcala's argument in the introduction to Section B.1 describes these attractions as "amenities" and therefore income distribution appears a reasonable choice for the attraction mass.

$$G_{PEi} = ID_i \quad (B.7)$$

$$A_{PEj} = ID_j \quad (B.8)$$

### B.1.3c Pleasure (PL)

This group poses a greater problem. The generating mass equals income

distribution again for the same reasons discussed for the previous two groups. The attraction mass, however, requires a different approach. The actual pleasure attractions vary: Hawaii offers beaches, Las Vegas offers gambling, and Orlando offers a fantasy world. Attempting to measure specific attractions directly is futile. An indirect approach suggests a measure like hotel rooms, but that number undoubtedly grows with a city's size. Of course if the number of hotel rooms (HR) was corrected for city size this measure might work well. For example, one could use hotel rooms divided by the total income as a possible variable. A pleasure oriented city would have a large ratio compared to other cities. An industrial or government center might have a large ratio also, particularly a city such as Washington, D.C. Further refinements might also be developed.

$$G_{PLi} = ID_i \quad (B.9)$$

$$A_{PLj} = HR_j / I_j \quad (B.10)$$

$I_j$  = Total income in city j

$HR_j$  = Number of hotel rooms in city j

#### B.1.3d Business (B)

Business travel needs different mass variables from those described above. Both the ability to generate and attract business travelers depends upon the level of business activity. A measure such as value added to manufactured goods is a reasonable business indicator (see Blumer (5)). So both the generation and attraction masses equal the business indicator (B).



$$G_i = B_i \quad (B.11)$$

$$A_j = B_j \quad (B.12)$$

### B.1.3e Summary

In the above discussion, the gravitational attraction is refined by splitting it into four types of attraction and by direction, and then developing appropriate masses for each city and each travel type. However, the derivation is still incomplete. No mention is made toward combining the four types of traffic into a single variable or combining  $i \rightarrow j$  traffic with the  $j \rightarrow i$  traffic to calculate total  $ij$  traffic, nor is such an attempt to be made here. These problems require additional research before the generation-attraction model can successfully produce valid mass variables.

### B.2 Distance

In building a model, careful attention must be given to the interactions occurring because distance correlates with several explanatory phenomena in transportation, including the desire to travel (LOI), the impedance to travel, and the choice of mode. Often modelers develop a variable to explain one of the above phenomena, without realizing that because of the correlations due to distance, several of the phenomena are affecting the results. Fare is an example of this, since it often enters a model exogenously when the modeler is searching for the price elasticity of demand. Unfortunately, the result is usually the combined LOI-impedance, mode elasticity of demand, with fare buried deep within the impedance. To simplify the example, assume time is the only other

distance -correlated explanatory phenomenon. If one enters fare as an independent variable, then the elasticity of fare is in fact the combined elasticity of fare and time. What fraction contributes to fare and what part to time is uncertain. This probably explains why attempts to find the price elasticity of demand have produced such erratic results. This section discusses in detail the several phenomena correlated with distance.

### B.2.1 Impedance

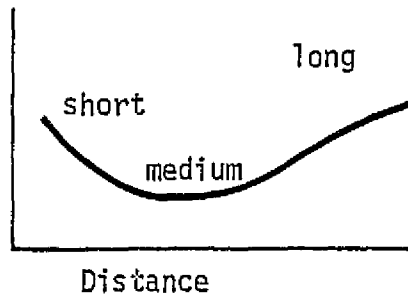
The cost of traveling, or the impedance, equals the trip time plus the fare. Trip time decomposes into waiting time and travel time, where waiting time equals the difference between the desired departure time and the actual departure time, and travel time equals the time spent en route on a mode (see Section B.3). The components of impedance are related to distance. The relationship, however, also depends upon city size, so the simple correlation between impedance and distance can be large or small.

Initially let us assume fixed city sizes, e.g., envision two cities, any two cities, connected by an elastic link, permitting the distance between them to change. As the link stretches the travel time and fare both increase directly. The waiting time changes also, but indirectly through a chain of events. The change in distance changes the demand due to LOI, modal-competition, and the other two impedance effects; in turn the change in demand affects the frequency of service; and finally the frequency change affects the waiting time. So given fixed city sizes, the impedance correlates strongly with distance (see Figures B.1 and B.2).

Now imagine a fixed distance, but the cities have elastic masses. As the masses grow the travel time and fare remain constant,

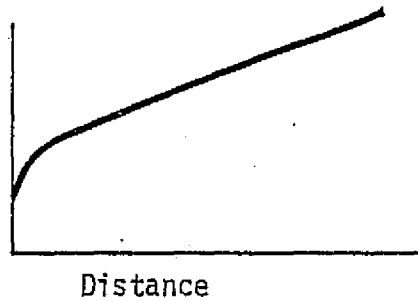
Figure B.1: The Components of Impedance Versus Distance

B.1a

Waiting  
Time

Waiting time in the short haul markets decreases at very short distances because demand increases (see Figure B.2)

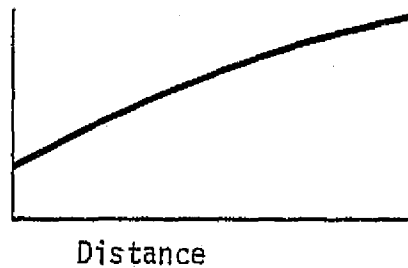
B.1b

Travel  
Time

For very short distances the take-off and landing impose a fixed time cost, but once into the air the time varies linearly with distance.

B.1c

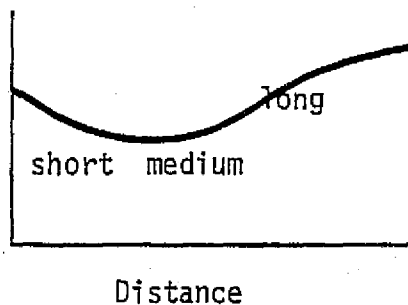
Fare



There is a fixed station charge and a per mile charge which tapers with distance

B.1d

Impedance

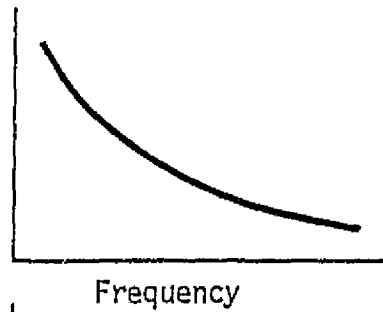


Sum curves B.1a through B.1c to yield B.1d

Figure B.2 Waiting Time as a Function of Inter-City Distance

B.2a

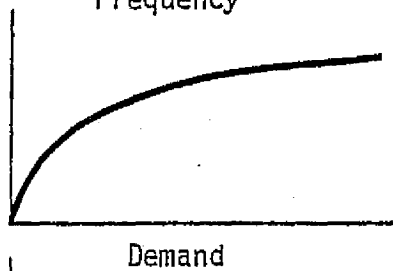
Waiting Time



Waiting time is a function of time between flights which decreases as frequency increases.

B.2b

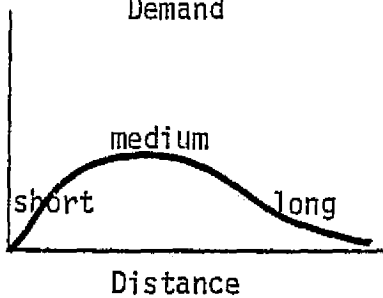
Frequency



The curve's decreasing slope reflects introduction of larger aircraft.

B.2c

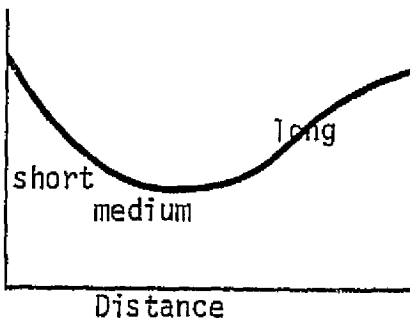
Demand



This curve is derived in Section B.2.2. It results from intermodal competition.

B.2d

Waiting Time



Combining curves B.2a-B.2c yields curve B.2d.

but the waiting time decreases because the larger masses cause higher traffic flows which in turn stimulate higher frequencies which mean lower waiting times.

For a sample of city-pairs varying in both distance and size, impedance does not correlate closely to distance. This results because of the variation in city sizes. Waiting time, a significant part of impedance for short hauls, is related not only to distance but to city masses. So for a given distance the impedance still varies widely because the city masses vary widely. In our sample, distance's correlation with air impedance was equal to .239 and with rail impedance equaled .268, both quite small.

However, should the waiting time equal zero for a mode, then no relationship to city size exists and the correlation between distance and impedance is quite high. For example, auto has no waiting time, and not surprisingly the auto impedance-distance correlation equaled .917 for our sample.

Therefore, scheduled modes show little correlation to distance because air is a scheduled mode, the impedance does not correlate with distance. Therefore, impedance should not cause interpretational problems because it will not be linked to other explanatory phenomena through distance. The harmonic mean of the modes has a modest correlation (.450), but not enough to cause great concern.

### B.2.2 Inter-Modal Competition

After deciding if he wants to travel and where he wants to travel, an individual must decide how he wants to travel, i.e., by what mode.

Within the abstract mode framework, the decision is based upon the mode attributes, trip time and cost. The short haul model for air travel, however, gives zero value to cost, so the decision reduces to comparing trip time or impedances.

For example, in comparing auto to air, auto demonstrates an advantage at very short distances, while air demonstrates an advantage at longer distances (see Figure B.3). Auto has no waiting time, a car sits in the driveway waiting to be driven at the owner's whim, while the public modes must meet published schedules. So an auto can travel miles down the road before an air passenger even steps onto the plane. The aircraft, however, is much faster than auto, and given a long enough distance to travel will pass the auto. So the auto has the shortest time for very short trips and aircraft has the shortest trip time for longer trips. One might ask what point are the two equal? This point varies from place to place but presumably lies near 100 miles.\*

Except for special cases, rail and bus never have a time advantage because their waiting time is as long as air and their speed is as slow as auto. Those two modes are for individuals without a car, or who do not care to drive, and have time but presumably not a lot of money. Also, those two modes may possess convenience characteristics in special circumstances.

Passenger generation by all modes and air's share of the traffic combine to produce a unique demand curve (Figure B.4). Total trips by all

---

\*Both the air and auto modes cover 100 miles in about two hours. For air assume 30 minutes to access airport and 30 minutes to egress airport. Then add 30 minutes waiting time in the terminal and 30 minutes for taxiing, flight, and landing. For auto assume an average speed of 50 mph.

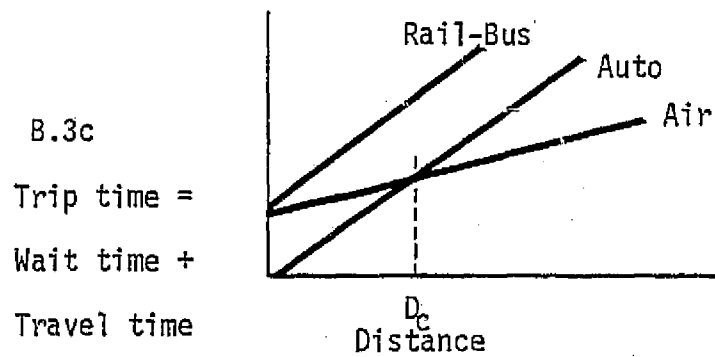
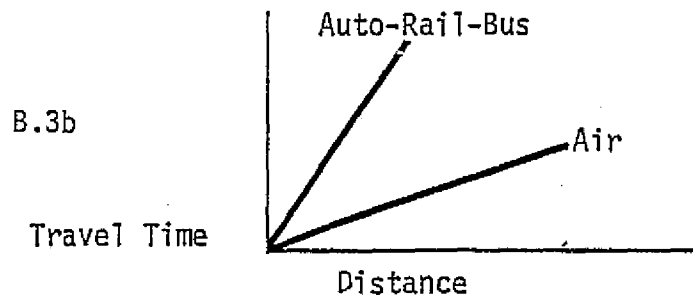
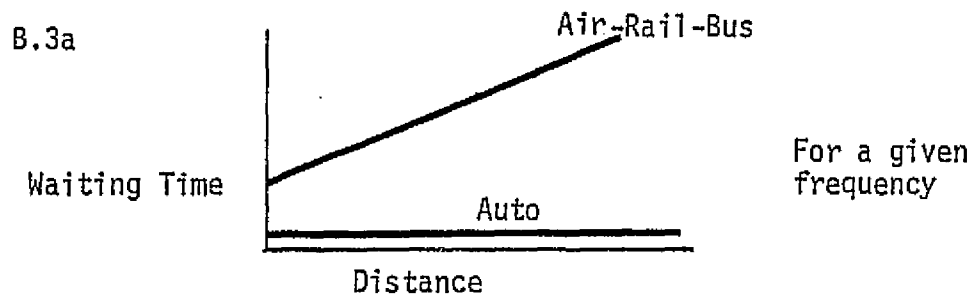
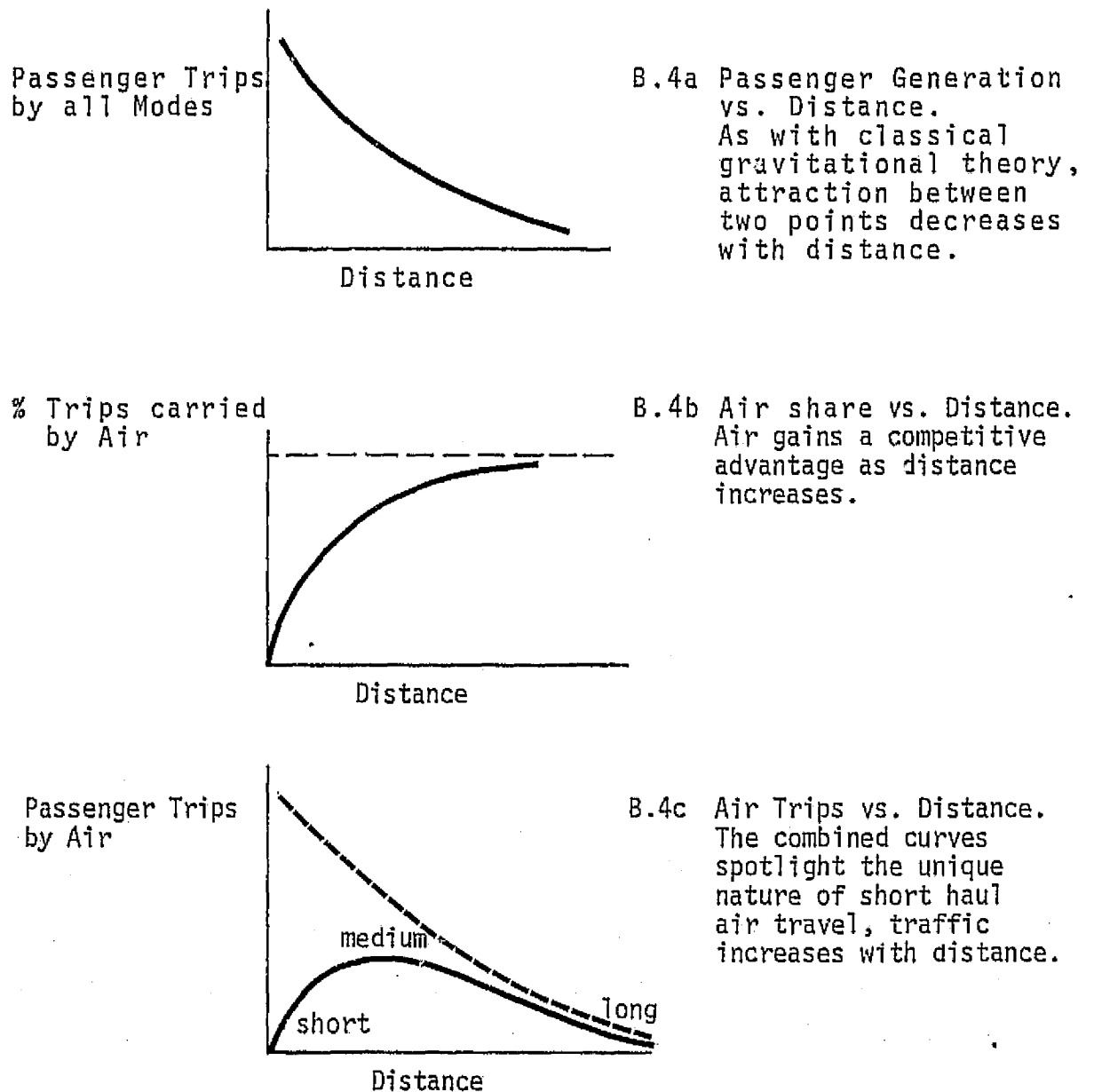
Figure B.3 Trip Times by Mode as a Function of Distance

Figure B.4 Air Traffic versus Distance



modes decrease with distance, due in part to level of interest and in part to impedance. Meanwhile air's share of the traffic steadily increases with distance as its competitive advantage increases. Combining these effects yields a passenger demand curve with a positive slope for the short haul and a negative slope for the medium and long hauls.

### B.2.3 Level of Interest

An individual normally has a reason for traveling; his destination holds an attraction which lures him from home. That attraction could be a business interest, a friend, or an event or activity of interest. But that attraction must exist before one desires to travel. So what is the probability that an attraction exists? For business interests and friends, distance is a strong measure of the probability. After all, is it not more likely one will make a friend next door than in the next town, and in the next town rather than in a city many miles away? This effect is called level of interest. The probability of being interested in a city decreases with distance.

Note the effect varies according to the type of attraction. One would suspect personal contacts have the highest correlation to distance, since these contacts develop with frequency of visits (contacts) with a given locale. Business contacts would be less correlated. Most small companies and local divisions of large companies do business locally, but corporate headquarters or main plants may be less correlated as companies span the globe. Finally, pleasure travel is probably the least correlated, as faraway places sound more exotic.

So level of interest is a very real effect for personal and business

travel, but has a questionable effect on pleasure travel. This model is business oriented due to its short haul nature, so the effect is important here. Note also that since long haul markets have a concentration of pleasure travelers, LOI has a small effect on demand in long haul markets. So LOI is not only defined by variations in distance, but coincidentally has a decreasing effect with distance. In other words, not only does the effect (LOI) decrease with distance, but the significance of the effect decreases with distance.

The impedance and level of interest effects may have a cause-effect relationship. Given two destinations, one with high travel impedance and the other without, a person will more likely develop personal or business contacts in the latter city. So impedance not only affects the decision to travel but it affects the degree of interest, and the degree of interest measures the desire to travel. So impedance as measured by time and cost may underlie LOI.

Distance correlates strongly with level of interest and inter-modal competition, but because our sample has cities of various sizes, distance does not correlate with impedance. The short haul model must therefore note the close connection between LOI and modes.

### B.3 Impedance

Demand models attempting to describe the service offered by a mode, usually do so by including a travel characteristic such as frequency, fare, or travel time. Unfortunately, none of these surrogates for service work particularly well, so more sophisticated variables have been developed. One example is UTRC's disutility or impedance, which equals the total cost of

a trip, both time and money costs. Impedance is used in the short haul demand model, but due to the special nature of short haul air travel, equals just the trip time.

The purpose of this discussion is to describe the derivation of trip time. The simplest measure equals travel time (block time for an aircraft). This measure ignores waiting time, however, and is therefore unable to reflect service accurately. For example, service with two departures a day is superior to one even if the travel times are equal. So an expected or average waiting time should also be calculated. This can be achieved by using a simple frequency function such as

$$\text{Waiting Time} = \left( \frac{\text{Hours in Traveling Day}}{\text{Flights per Day}} \right) * 1/2$$

For example, if a traveling day equals 16 hours and 4 flights are available, the average time between flights equals 4 hours and the expected waiting time equals 2 hours. This measure is still very crude, however, and completely ignores the behavioral nature of a traveler.

Steve Eriksen has investigated this problem in depth, and has proposed a behavioral model which describes flight selection and calculates expected waiting time.(22) In Eriksen's model, the "preferred departure time" or PDT model assumes an individual has a desired departure time, but due to airline scheduling must leave earlier or later than desired. The flight he selects is that flight which minimizes his trip time. His trip time equals the sum of his travel time and displacement (waiting) time, where his displacement time equals the absolute value of difference between his desired and actual departure times.

Several important assumptions are embodied in the model. First, all times are weighted equally. Waiting time, travel time, and time lost during a stop are all viewed the same.\*

In addition the passenger demand is assigned to flights without regard to capacity, i.e. infinite capacity is assumed. Finally, Eriksen derives demand distributions that are far more realistic than a uniform time-of-day distribution. The model is by no means perfect, but is far better than any other built to date, and therefore is used here.

Though designed with air transportation in mind, the model works equally well for rail, in fact any mode. Auto, of course, has zero waiting time and therefore is a degenerate case.

#### B.4 Fare

Besides a time cost, impedance typically includes a money cost represented by the fare for a public mode and by fuel, toll, meal, and lodging costs for auto. Looking just at public modes for a moment, one finds many options exist for constructing a fare variable because a number of fares are available, e.g. first class, coach, and discount. Typically, modelers use coach fare or some weighted average. This short haul model uses the standard coach fare. Besides being relatively easy to collect, research performed at M.I.T. and by Verleger indicates coach to be the proper fare to use, and at least is no worse than any other option. (19,41)

Calibrations on a simple demand model at M.I.T. indicate that any consistent fare variable will work equally well. The study investigated three variables: coach, estimated average fare, and actual average fare, and found

---

\* Eriksen may change this assumption in refined versions of the model.

no significant difference to exist between the models. As long as no significant difference exists, the coach fare appears most attractive because of its simplicity, i.e., it requires no special calculations or massive data collection efforts.

Comparing the results expressed in equations (5.10-12) indicates that the variable coefficients, the  $t$  ratios, and the coefficients of multiple determination ( $R^2$ ) do not vary significantly between the models. The conclusion drawn from this analysis is that the respective elasticities, their precisions, and the prediction accuracy of the models are independent of the fare variable selected. Therefore, any reasonable fare variable used in such a model should produce equivalent results (19).

Verleger attacks the question on theoretical level. He notes that the true demand curve results from the fare which attracts the marginal passenger. That passenger is attracted by the lowest available fare, not some weighted average. If one uses a weighted average the price elasticity will be greater or equal to the true value. The only time the weighted average yields the true elasticity is when all fares change by the same percentage, as apparently occurred in the data for the M.I.T. study. At any rate, the theoretically correct fare is the lowest available fare, a case also included in the M.I.T. study.

Next one must determine what the lowest available fare is, a more complicated problem than one might think. Both coach and discount fares are candidates. The discount fare is lower but is not available to everyone, most discount fares require 14 day advance ticketing, 7 day minimum stay, and travel on pre-specified days only. As noted earlier, most short haul travelers are businessmen to whom time is very important. They cannot commit themselves two weeks in advance for a one-day trip; for that matter

they cannot stay 7 days for a one-day trip. So one concludes that discount fares are impractical for most short haul travelers. Therefore the lowest available fare is coach. So after inspecting M.I.T.'s study and Verleger's work, the coach fare was selected for the short haul model.

The rail fare was set equal to the coach fare for the same reasons as given for air. However, one more note must be made. Rail attracts many fare-conscious individuals even in the short haul markets. These travelers do not affect this study's choice of coach fare, because the travelers of interest in this study are just those who are likely to travel by air in short haul markets. These individuals comprise only a fraction of the total rail demand.

Auto does not have a fare in the conventional sense, but does have costs all the same. A traveler must purchase fuel, pay tolls, and cover meal and lodging expenses. This study uses a cost equation developed by UTRC which appears to be quite reasonable.<sup>(34)</sup> The hourly cost covers meals and lodging while the mileage cost covers fuel, oil and tolls. Depreciation and insurance do not appear because both are incurred whether or not the trip is taken.

$$\text{Cost (\$)} = 20(\$/\text{HR}) * \text{Travel Time (HRS)} + .051(\$/\text{MI}) * \text{Distance (MI)}$$

(B.13)

## APPENDIX C

## List of Tables Pertaining to Model Results

This appendix includes 15 tables that depict the results of the base model, the mode sensitive model, the destination sensitive model, and the combined mode-destination sensitive model. The discussion of the development of these models appears in PART I of this report.

Table C.1: Search for Optimal  $a$  and  $V$ .\*

$a$	$V$	$R^2$	$F$	$b_0$	$b_1$	$t_1$	$b_2$	$t_2$
-2	.2	.28	26	12.1	.40	5.1	-.33	.7
-2	.1	.32	33	12.4	.28	3.7	-1.30	3.1
-1.5	.2	.34	36	11.3	.50	6.6	-.21	-.5
-1.5	.1	.38	42	11.9	.39	5.2	-1.05	-2.8
-1	.2	.42	49	10.2	.57	8.2	-.24	-.6
-1	.1	.45	55	10.9	.49	6.9	-.93	-2.7
-1	0	.57	88	10.1	.32	4.9	-1.40	-6.9
-1.2	.09	.43	51	11.4	.44	6.0	-1.40	-3.0
-1.2	.08	.43	52	11.4	.42	5.9	-1.12	-3.3
-1.2	.05	.47	59	11.2	.38	5.3	-1.30	-4.4
-.8	.09	.49	64	10.4	.51	7.4	-.97	-3.0
-.8	.08	.49	65	10.4	.50	7.3	-1.04	-3.3
-.8	.05	.52	72	10.4	.45	6.7	-1.21	-4.3
-.5	.09	.53	76	9.6	.55	8.5	-.97	-3.2
-.5	.08	.53	77	9.1	.54	8.4	-1.02	-3.5
-.5	.05	.56	85	8.9	.50	7.8	-1.17	-4.5
-.4	.04	.58	93	9.3	.50	8.0	-1.2	-4.9
-.4	.02	.60	102	9.1	.47	7.6	-1.3	-5.8
-.4	0	.63	116	8.9	.43	7.1	-1.3	-6.8
-.2	.04	.60	103	8.7	.52	8.8	-1.2	-5.1
-.2	.02	.63	113	8.6	.49	8.3	-1.2	-5.9
-.2	0	.65	127	8.4	.46	7.9	-1.2	-6.9
0	.04	.63	114	8.1	.54	9.5	-1.2	-5.4



Table C.1 (Continued)

a	V	R <sup>2</sup>	F	b <sub>0</sub>	b <sub>1</sub>	t <sub>1</sub>	b <sub>2</sub>	t <sub>2</sub>
0	.02	.65	124	8.0	.51	9.1	-1.2	-6.1
0	0	.67	139	7.8	.48	8.6	-1.2	-7.1
.5	0	.72	169	6.4	.51	10.3	-1.21	-7.9
1	0	.75	203	5.0	.51	11.9	-1.26	-9.0
1.5	0	.78	233	3.9	.49	13.2	-1.33	-10.4
2	0	.79	256	3.0	.47	14.0	-1.42	-11.7
2.5	0	.80	269	2.3	.43	14.5	-1.51	-12.9
3	0	.80	276	1.8	.40	14.7	-1.59	-13.9

\* Calibrated on equation (2.1)

$$T_{ij} = b_0 G_{ij} + b_1 I_{ija} + b_2$$

Table G.2 Proxies for Mode Sensitive Model

Simple Type Proxies

- 1)  $D$  = Distance
- 2)  $I_u$  = Impedance fo auto
- 3)  $H$  = Impedance of modes other than air (the harmonic mean of auto and rail)
- 4)  $\bar{I}$  = Impedance of all modes (the harmonic mean of air, auto and rail)

Ratio Type Proxies

- 5)  $I_u/I_a$  = The impedance of air relative to the impedance
- 6)  $H/I_a$  = The impedance of air relative to the impedance of auto and rail ("other" modes)
- 7)  $\bar{I}/I_a$  = The impedance of air relative to the impedance of auto, rail and air (all modes)

Difference Type Proxies

- 8)\*  $(I_u - I_a + 10)/D = DA$  = The impedance of air relative to the impedance of auto
- 9)\*  $(H - I_a + 10)/D = DIFF$  = The impedance of air relative to the auto and rail

---

\*The constant 10 insures that the proxy never equals zero. This condition is necessary because the logarithm of zero equals minus infinity.

Table 0.3 Air share Variable ( $f_{ij}$ )Simple Type Variables

- 1)  $D =$  Distance

Ratio Type Variables

- 2)  $\bar{I}/I_a =$  The impedance of air relative to the impedance of auto, rail, and air
- 3)  $H/I_a =$  The impedance of air relative to the impedance of auto and rail
- 4)  $I_u/I_a =$  The impedance of air relative to the impedance of auto

Difference Type Variables

- 5)  $(H - I_a + 10)/D = \text{DIFF} =$  The impedance of air relative to the impedance of auto and rail
- 6)  $(I_u - I_a + 10)/D = \text{DA} =$  The impedance of air relative to the impedance

Table C.4: Results for Proxy Specification  
of Mode Sensitive Model

	$R_2$	F	$b_0$	$b_1$	$t_1$	$b_2$	$t_2$	$b_3$	$t_3$
$G * I_a$	.57	88	10.1	.32	4.9	-1.4	-6.9		
$G * I_a \times D$	.81	183	1.5	.38	8.7	-1.64	-11.8	1.65	12.7
$G * I_a \times I_u$	.79	162	7.8	.31	6.7	-1.74	-11.8	1.50	11.6
$G * I_a \times H$	.75	136	8.2	.35	7.2	-1.74	-11.0	1.36	10.1
$G * I_a \times \bar{I}$	.68	94	10.2	.34	6.1	-4.57	-9.3	3.88	6.9
$G * I_a \times I_u/I_a$	.79	162	7.7	.31	6.7	-.24	-1.4	1.50	11.6
$G * I_a \times H/I_a$	.75	134	7.9	.37	7.4	-.38	-2.0	1.36	10.0
$G * I_a \times \bar{I}/I_a$	.68	94	10.2	.34	6.1	-.68	-3.3	3.88	10.0
$G * I_a$ DIFF	.81	185	3.7	.34	8.0	-2.08	-14.2	-2.58	-12.9
$G * I_a$ DA	.81	189	3.7	.38	8.8	-2.19	-14.7	-2.58	-13.0

Table 0.5: Correlations for Mode Sensitive Model

	D	$I_u$	H	$\bar{I}$	$I_u/I_a$	$H/I_a$	$\bar{I}/I_a$
$I_a$	.239	.227	.321	.957	-.661	-.601	-.572
$\bar{I}$	.450	.446	.532	-	-.457	-.385	-.310
M/D	-.23	-.122	-.267	-.591	.410	.299	.302
TRIPS	.273	.301	.156	-.577	.817	.741	.661
D	-	.917	.891	.450	.506	.544	.486

	DIF	DA	$I_a$	M/D	TRIPS
$I_a$	-.537	-.728	-	-.602	-.699
$\bar{I}$	-.563	-.671	.957	-.591	-.577
M/D	.314	.401	-.602	-	.644
TRIPS	-.010	-.101	-.699	.644	-
D	-.898	-.903	.239	-.23	.273

Table C.6: Simple and Partial Correlations of Distance (D) and Difference on Auto (DA) with the Dependent Variable TRIPS (T).

Correlaiton Between	Also Included in the Model	Correlation Coefficients	
		Simple	Partial
D and TRIPS		.273	
D and TRIPS	I <sub>a</sub>		.635
D and TRIPS	I <sub>a</sub> and G		.744
DA and TRIPS		-.010	
DA and TRIPS	I <sub>a</sub>		-.640
DA and TRIPS	I <sub>a</sub> and G		-.749

Table C.7 Air Share Model Test Results

	$R_2$	F	$b_0$	$b_1$	$t_1$	$b_2$	$t_2$	$b_3$	$t_3$
BASE MODEL	.59	88	10.1	.32	4.9	-1.4	-6.9		
$G * \bar{I} * D$	.81	187	-.9	.39	9.8	-2.05	-12.1	2.12	15.3
$G * \bar{I} * I_u/I_a$	.79	164	7.7	.30	6.5	-.30	- 1.7	1.55	14.0
$G * \bar{I} * H/I_a$	.75	135	7.8	.37	7.5	-.39	- 2.1	1.45	12.3
$G * \bar{I} * \bar{I}/I_a$	.68	94	10.2	.34	5.7	-.68	- 3.3	4.57	9.3
$G * \bar{I} * DA$	.76	144	2.1	.43	8.6	-2.63	-12.0	-3.14	-12.6
$G * \bar{I} * DIFF$	.75	134	2.4	.39	8.0	-2.37	-11.2	-3.06	-12.2

Table C.8 Three Best Specificaitons of the  
Mode Sensitive Model

	$R_2$	F	$b_0$	$b_1$	$t_1$	$b_2$	$t_2$	$b_3$	$t_3$
Base $G * I_a$	.59	88	10.1	.32	4.9	-1.4	-6.9		
$G * I_a * D$	.81	183	1.5	.38	9.5	-1.64	-11.7	1.65	12.7
$G * I_a * DA$	.81	189	3.7	.38	9.5	-2.19	-14.7	-2.54	-12.7
$G * \bar{I} * D$	.81	187	-.9	.139	9.8	-2.05	12.1	2.12	15.1



Table C.9: Comparison of Exponents Used to  
Derive the Harmonic Mean

Exponent	Model	$R_2$	F	$b_0$	$b_1$	$t_1$	$b_2$	$t_2$	$b_3$	$t_3$
2	$G * \bar{I} * I_u/I_a$	.79	164	7.7	.30	6.5	-.30	-1.7	1.55	14.0
1.5	$G * \bar{I} * I_u/I_a$	.79	164	7.7	.30	6.4	-.31	-1.8	1.56	14.4
1	$G * \bar{I} * I/I_a$	.79	165	7.6	.29	6.3	-.34	-2.0	1.57	14.9
2	$G * \bar{I} * H/I_a$	.75	135	7.8	.37	7.5	-.39	-2.1	1.45	12.3
1.5	$G * \bar{I} * H/I_a$	.74	127	8.2	.36	7.3	-.46	-2.5	1.40	12.1
1	$G * \bar{I} * H/I_a$	.71	111	8.4	.37	7.1	-.59	-3.1	1.35	11.4

Table 0.10: Destination Sensitive Results

Model	$R^2$	$\sigma$	F	$b_0$	$B_1$	$t_1$	$b_2$	$t_2$	$b_3$	$t_2$
<u>Source</u>										
$b_1 = (M \ S)^{\theta_1}$					$H_{ij1}$		$I_{ija}$		$F_{ij1}$	
$\theta_1 = .4$	.59	.96	64	10.1	.85	5.2	-1.06	-4.5	.65	4.7
.5	.59	.96	64	10.2	.67	5.2	-1.07	-4.5	.64	4.6
.6	.59	.96	64	10.3	.55	5.1	-1.08	-4.5	.62	4.5
1.0	.58	.98	63	10.3	.31	4.8	-1.21	-5.1	.50	4.0
$b_2 = M^{\theta_2}$										
					$H_{ij2}$		$I_{ija}$		$F_{ij2}$	
$\theta_2 = .8$	.71	.81	107	7.4	1.25	9.6	-.79	-4.0	.36	3.6
.9	.71	.81	106	7.5	1.11	9.6	-.80	-4.0	.36	3.7
1.0	.71	.81	106	7.6	.90	9.6	-.80	-4.1	.35	3.8
1.1	.70	.81	106	7.6	.99	9.6	-.80	-4.1	.36	3.8
<u>Proxy</u>										
					$G_{ij}$		$I_{ija}$		S	
S + S	.65		81	10.1	.63	7.9	-.81	3.7	-.62	-5.6
S * S	.65		83	9.7	.65	8.1	-.78	3.5	-.31	-5.2
SW + SW	.64		79	9.8	.59	7.4	-.84	3.8	-.56	-5.1
BASE MODEL	.59		88	10.1	.32	4.9	-1.4	-6.9		

Table C.11: Comparison of values of the distance exponent in calculation of S

$$T_{ij} = b_0 (M_i * M_j * D^{1.5})^{b_1} I_{ija}^{b_2} (S_i + S_j)^{b_3}$$

Exponent	R <sup>2</sup>	F	b <sub>0</sub>	b <sub>1</sub>	t <sub>1</sub>	b <sub>2</sub>	t <sub>2</sub>	b <sub>3</sub>	t <sub>3</sub>
0	.84	237	5.05	.60	17.2	-1.17	-10.6	-.51	-7.4
.3	.84	238	4.15	.60	17.3	-1.17	-10.7	-.49	-7.5
.7	.84	240	3.04	.60	17.4	-1.18	-10.7	-.47	-7.6
1.0	.84	241	2.3	.61	17.4	-1.18	-10.7	-.45	-7.6
1.3	.85	243	1.6	.61	17.5	-1.18	-10.8	-.45	-7.7
1.5	.84	240	3.1	.61	17.5	-1.18	-10.8	-.40	-7.7
1.6	.84	240	2.9	.61	17.5	-1.18	-10.8	-.40	-7.7
1.7	.84	240	2.7	.61	17.5	-1.18	-10.8	-.39	-7.7
1.8	.84	240	2.5	.61	17.5	-1.18	-10.8	-.38	-7.7
1.9	.84	239	2.3	.60	17.5	-1.18	-10.8	-.36	-7.7
2.0	.84	242	2.1	.61	17.5	-1.18	-10.8	-.36	-7.7

Table C.12: Range of Coefficients for  $I_a$  and D

Variable	Range of Coefficient
$I_a$	-1.19 to -1.75
D	1.33 to 1.87

Table C.13: Results for the Mode-Destination Sensitive  
Specifications of the Short Haul Demand Model

Model	$R^2$	F	$b_0$	$b_1$	$t_1$	$b_2$	$t_2$	$b_3$	$t_3$	$b_4$	$t_4$
G * I	.59	88	10.1	.32	4.9	-1.4	-6.9				
G * $I_a$ * D * S	.84	179	2.2	.60	11.0	-1.19	-8.1	153	13.0	-.44	-5.8
G * $I_a$ * DA * S	.84	173	4.3	.58	10.3	-1.75	-10.1	-2.33	-12.6	-.39	-4.9
G * $\bar{I}$ * D * S	.84	179	.4	.60	11.0	-1.50	-8.1	187	14.1	-.43	-5.5
H * $I_a$ * D * F	.86	197	1.1	.85	11.5	-1.19	-8.4	1.33	11.8	.37	5.5
H * $I_a$ * DA * F	.86	197	2.8	.85	11.5	-1.66	-10.6	-2.02	-11.8	.34	5.1
H * $\bar{I}$ * D * F	.86	195	-.7	.86	11.8	-1.49	-8.3	1.65	12.9	.36	5.3

Table C.14: Comparing Coefficients Across  
Three Years.

Specifi- cation	Coeffi- cient	1972-74	1972	1973	1974
1	$b_1$	.60	.6990	.7125	.4103
	$b_2$	-1.19	-.9447	-.8969	-1.6488
	$b_3$	1.53	1.3882	1.6448	1.5237
	$b_4$	-.44	-.5337	-.5299	-.2392
2	$b_1$	.58	.6996	.6427	.3929
	$b_2$	-1.75	-1.3410	-1.6070	-2.2549
	$b_3$	-2.33	-1.9529	-2.5273	-2.5048
	$b_4$	-.39	-.5316	-.4201	-.1638
3	$b_1$	.60	.7263	.6931	.4152
	$b_2$	-1.50	-1.3410	-1.2091	-2.2009
	$b_3$	1.87	-1.5645	-.4958	2.1316
	$b_4$	-.43	-.5499	.4378	-.2187
4	$b_1$	.85	.9341	.9672	.6702
	$b_2$	-1.19	-.9444	-.9534	-1.5576
	$b_3$	1.33	1.2198	1.4159	1.2856
	$b_4$	.37	.5073	.4378	.1743
5	$b_1$	.85	.9475	.9309	.6976
	$b_2$	-1.66	-1.0398	-1.5316	-2.0326
	$b_3$	-2.02	-2.7035	-2.1657	-2.1143
	$b_4$	.37	.5015	.3864	.1525
6	$b_1$	.86	.9677	.9559	.6908
	$b_2$	-1.49	-1.0398	-1.2554	-2.0797
	$b_3$	1.65	1.3867	1.6970	1.8404
	$b_4$	.36	.5300	.4194	.1549

Table C.15: Results for models calibrated on 1972 and 1973 data only

Model	$R^2$	$\sigma$	F	$b_0$	$b_1$	$t_1$	$b_2$	$t_2$	$b_3$	$t_3$	$b_4$	$t_4$
G * $I_a$ * D * S	.88	.5254	167	2.2	.70	11.9	-.93	-5.8	1.52	11.8	-.53	-6.6
G * $I_a$ * DA * S	.87	.5593	145	4.5	.68	10.7	-1.46	-7.7	-2.22	-10.7	-.49	-5.6
G * $\bar{I}$ * D * S	.88	.5340	162	1.0	.71	11.9	-1.11	-5.5	1.74	12.0	-.53	-6.4
H * $I_a$ * D * F	.90	.4984	188	.9	.95	12.3	-.95	-6.2	1.32	11.0	.47	6.6
H * $I_a$ * DA * F	.88	.5254	169	2.9	.94	11.6	-1.40	-7.8	-1.93	-10.1	.45	5.9
H * $\bar{I}$ * D * F	.89	.5075	180	-.3	.96	12.3	-1.13	-5.7	1.54	11.2	.48	6.5